

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:19:16 : Search time 86 Seconds
(without alignments)
804.707 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MKRPYVAIVGRPNVKSTIF.....IRAFGEFGPIHIARKRN 436

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2213	100.0	436	AAU37007	Staphylococcus aur
2	2213	100.0	436	AAU37139	Staphylococcus aur
3	2213	100.0	436	AAU37139	Staphylococcus aur
4	2207	99.7	435	AAU37139	Staphylococcus aur
5	2094	94.6	443	ABP38338	Staphylococcus aur
6	1731	78.2	436	ABP38338	Staphylococcus epi
7	1528	69.0	436	ABP27827	Streptococcus poly
8	1524	68.9	436	ABP27827	S. epidermidis ope
9	1523	68.8	436	ABP27826	Streptococcus poly

10	1522	68.8	436	21	AAU04108	lyphc protein of st
11	1522	68.8	436	21	AAU04108	Streptococcus pneu
12	1522	68.8	436	24	ABU02185	S. pneumoniae type
13	1520	68.7	436	23	ABU04058	Lactococcus lactis
14	993.5	44.9	448	22	AAU09939	ERA binding domain
15	935.5	42.3	448	22	AAU09941	ERA binding domain
16	923	41.7	503	22	AAU04658	E. coli cellular p
17	907	41.0	501	22	AAU09942	ERA binding domain
18	899	40.6	490	22	AAU08110	Salmonella typhi c
19	862.5	39.0	519	22	AAU03123	C glutamic prote
20	857.5	38.7	500	22	ABP78407	N. gonorrhoeae ami
21	852	38.5	504	22	AAU035376	Haemophilus influe
22	852	38.5	504	22	AAU08526	Haemophilus influe
23	852	38.5	504	23	AAU01457	Haemophilus influe
24	851.5	38.5	709	23	ABP65695	Bifidobacterium lo
25	845.5	38.2	461	22	AAU00718	Propionibacterium
26	841.5	38.0	246	22	AAU07925	Streptococcus pneu
27	840.5	38.0	493	22	AAU06363	Pseudomonas aerugi
28	835	37.7	502	22	AAU09943	ERA binding domain
29	799.5	36.1	459	22	AAU09946	ERA binding domain
30	776.5	35.1	459	22	AAU09945	ERA binding domain
31	754.5	34.1	225	22	AAU033525	Enterococcus faeca
32	750.5	33.9	225	22	AAU03506	Enterococcus faeca
33	708.5	32.0	493	20	AAU07461	Amino acid sequenc
34	707	31.9	456	22	AAU09937	ERA binding domain
35	706	31.9	458	22	AAU05777	Helicobacter pylor
36	685.5	31.0	487	20	AAU05489	Amino acid sequenc
37	669.5	30.3	453	22	AAU09944	ERA binding domain
38	624	28.2	433	22	AAU09948	ERA binding domain
39	620.5	28.0	530	22	AAU09947	ERA binding domain
40	613	27.7	292	22	AAU09940	ERA binding domain
41	559.5	25.3	383	17	AAU07678	Campylobacter jelu
42	528.5	23.9	331	23	ABU01840	Helicobacter pylor
43	433.5	19.6	208	22	AAU09938	ERA binding domain
44	340	15.4	106	22	AAU09939	ERA binding domain
45	315	14.2	92	19	AAU06046	S. pneumoniae deri

ALIGNMENTS

RESULT 1	AAU37007	standard; Protein: 436 AA.
ID	AAU37007	
XX	AAU37007:	
AC	14-FEB-2002 (first entry)	
XX		
DT	Staphylococcus aureus cellular proliferation protein #1177.	
XX		
DE	Antisense: prokaryotic cellular proliferation protein:	
XX	antibiotic; antibacterial; drug design.	
KW	Staphylococcus aureus.	
XX		
OS	W0200170955-A2.	
XX		
PN	27-SEP-2001.	
XX		
PD	21-MAR-2001; 2001WO-US09180.	
XX		
PF	21-MAR-2000; 2000US-206848P.	
XX		
PR	26-MAY-2000; 2000US-207727P.	
XX		
PR	23-OCT-2000; 2000US-242578P.	
XX		
PR	27-NOV-2000; 2000US-253625P.	
XX		
PR	22-DEC-2000; 2000US-257931P.	
XX		
PR	16-FEB-2001; 2001US-269308P.	
XX		
PA	(ELIT-) ELITRA PHARM INC.	
XX		
PI	Hasselbeck R, Ohlsen KL, Zysek JD, Wall D, Trawick JD, Carr GJ;	

PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR N-PSDB: AAS54866.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12600; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pctl_sequences.
XX
SQ Sequence 436 AA:
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4,7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERYSIYEDTPGVTNRDRYSSGEWLTDFDNITIDGG 60
DB 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERYSIYEDTPGVTNRDRYSSGEWLTDFDNITIDGG 60
QY 61 IEIGDAPFQTOIRAQAETAIDEADVIIFMVNVRREGLTQSDENVVAQILKSKKPPVLA VNK 120
DB 61 IEIGDAPFQTOIRAQAETAIDEADVIIFMVNVRREGLTQSDENVVAQILKSKKPPVLA VNK 120
QY 121 VDNMEMRTDYVDFYSLGGEPPYISGSHGLGLDLDAVVSHFGEEDDPYDEDTIRLSI 180
DB 121 VDNMEMRTDYVDFYSLGGEPPYISGSHGLGLDLDAVVSHFGEEDDPYDEDTIRLSI 180
QY 181 IGRNVNGSSLVNAVLGDRIVSVNAGTTRDAIDTEYSYGQDYVLIDTAGMRKKGVY 240
DB 181 IGRNVNGSSLVNAVLGDRIVSVNAGTTRDAIDTEYSYGQDYVLIDTAGMRKKGVY 240
QY 241 ESTEKYSVLRALKAIERSNVVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKM DIVE 300
DB 241 ESTEKYSVLRALKAIERSNVVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKM DIVE 300
QY 301 KDSKTMKKFEDEVKKEFOFLDYAQIAFVSAKERTRLRTLPYIN EASENHRKRVOSSTLN 360
DB 301 KDSKTMKKFEDEVKKEFOFLDYAQIAFVSAKERTRLRTLPYIN EASENHRKRVOSSTLN 360
QY 361 EYVVDATSMNPTPPDKGRLLNVFAATQVAIEPPTFFVNVNDELHMFYSKRYLENOI TAA 420
DB 361 EYVVDATSMNPTPPDKGRLLNVFAATQVAIEPPTFFVNVNDELHMFYSKRYLENOI TAA 420
QY 421 FGEGGTPIHITARRKN 436
DB 421 FGEGGTPIHITARRKN 436

RESULT 2
AAU37139
ID AAU37139 standard; Protein: 436 AA.
XX

AC AAU37139;
XX
XX 14-FEB-2002 (first entry)
DT
XX Staphylococcus aureus cellular proliferation protein #1309.
DE
XX Antisense: prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX
XX WO200170955-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-0509180.
PR
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI: 2001-611495/70.
DR N-PSDB: AAS54998.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
XX Example 3; Seq ID No 12732; 511pp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pctl_sequences.
XX
SQ Sequence 436 AA:
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4,7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERYSIYEDTPGVTNRDRYSSGEWLTDFDNITIDGG 60
DB 1 MTKPIVAIVGRPNVNGKSTIFNRIVGERYSIYEDTPGVTNRDRYSSGEWLTDFDNITIDGG 60
QY 61 IEIGDAPFQTOIRAQAETAIDEADVIIFMVNVRREGLTQSDENVVAQILKSKKPPVLA VNK 120
DB 61 IEIGDAPFQTOIRAQAETAIDEADVIIFMVNVRREGLTQSDENVVAQILKSKKPPVLA VNK 120
QY 121 VDNMEMRTDYVDFYSLGGEPPYISGSHGLGLDLDAVVSHFGEEDDPYDEDTIRLSI 180

|||||
Db 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDAAVSHFGEEEDPYDEDTRLISI 180
QY 181 IGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 240
Db 181 IGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 240
QY 241 ESTEKYSVLRAKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGOKAVIIVNKMWDIVE 300
Db 241 ESTEKYSVLRAKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGOKAVIIVNKMWDIVE 300
QY 301 KDSKTKMKKFEDEVRKREFOLDYAQIAFVSAKERTRLTLFPYINEASENHKKRVQSTLN 360
Db 301 KDSKTKMKKFEDEVRKREFOLDYAQIAFVSAKERTRLTLFPYINEASENHKKRVQSTLN 360
QY 361 EYVYDAISMNPPTDQGRRLNFFVATQVAIEPPTFVFVNDVLMHFSYKRYLENDQIRAA 420
Db 361 EYVYDAISMNPPTDQGRRLNFFVATQVAIEPPTFVFVNDVLMHFSYKRYLENDQIRAA 420
QY 421 FGFEETPIHIIARKRN 436
Db 421 FGFEETPIHIIARKRN 436

RESULT 3
AAB82089
ID AAB82089 standard; Protein: 436 AA.

XX AAB82089;

DT 26-JUN-2001 (first entry)

Staphylococcus aureus Yphc protein sequence.

Yphc: antimicrobial; cytostatic; antiulcer; microbial infection;
gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
gastritis.

Staphylococcus aureus.

WO200123418-A1.

05-APR-2001.

19-SEP-2000; 2000WO-US25566.

28-SEP-1999; 99US-0406968.

(SMIK) SMITHKLINE BEECHAM CORP.
(SMIK) SMITHKLINE BEECHAM PLC.

Zalacain M, Biswas S, Burnham MKR, Sylvester D, Mcdevitt D;
Matile TB;

WPI; 2001-308138/32.
N-PSDB; AAF86461.

Novel yphc polypeptides of Staphylococcus aureus useful for diagnosing
and treating microbial infections, especially infection by
Staphylococcus aureus and Helicobacter pylori

Claim 1; Page 3; 41pp; English.

The present sequence is the yphc protein of Staphylococcus aureus. The
yphc coding sequence and protein are useful for treating and diagnosing
microbial infections such as infection caused by S.aureus and
Helicobacter pylori. In addition, the yphc coding sequence and protein
are useful for treating diseases such as H.pylori-induced cancers, e.g.
gastrointestinal carcinoma, gastric ulcers, and gastritis. The present
sequence was obtained from a library of clones of chromosomal DNA of
S.aureus in E.coli. The sequencing data from two or more clones
comprising overlapping S.aureus DNAs was used to construct the present
contiguous DNA sequence.

XX SQ Sequence 436 AA;
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4.7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVAIYGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 60
Db 1 MTKPIVAIYGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 60
QY 61 IEIGDAPFQTOIRAQAEIADIDADVIFVNVNREGLTQSDENVAQILYKSKRPVLAUNK 120
Db 61 IEIGDAPFQTOIRAQAEIADIDADVIFVNVNREGLTQSDENVAQILYKSKRPVLAUNK 120
QY 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDAAVSHFGEEEDPYDEDTRLISI 180
Db 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDAAVSHFGEEEDPYDEDTRLISI 180
QY 181 IGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 240
Db 181 IGRPNVGSLSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGKY 240
QY 241 ESTEKYSVLRAKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGOKAVIIVNKMWDIVE 300
Db 241 ESTEKYSVLRAKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGOKAVIIVNKMWDIVE 300
QY 301 KDSKTKMKKFEDEVRKREFOLDYAQIAFVSAKERTRLTLFPYINEASENHKKRVQSTLN 360
Db 301 KDSKTKMKKFEDEVRKREFOLDYAQIAFVSAKERTRLTLFPYINEASENHKKRVQSTLN 360
QY 361 EYVYDAISMNPPTDQGRRLNFFVATQVAIEPPTFVFVNDVLMHFSYKRYLENDQIRAA 420
Db 361 EYVYDAISMNPPTDQGRRLNFFVATQVAIEPPTFVFVNDVLMHFSYKRYLENDQIRAA 420
QY 421 FGFEETPIHIIARKRN 436
Db 421 FGFEETPIHIIARKRN 436
RESULT 4
AAU33787
ID AAU33787 standard; Protein: 435 AA.
XX AAU33787;
DT 14-FEB-2002 (first entry)
DE Staphylococcus aureus cellular proliferation protein #63.
KW Antisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
OS
XX
PN WO200170955-A2.
PD 27-SEP-2001.
PF 21-MAR-2001; 2001WO-US09180.
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
PA
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.
DR N-PSDB; AAS51646.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS Example 3; Seq ID No 5283; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 435 AA:
Query Match 99.7%; Score 2207; DB 22; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.6e-183;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTTHDNIITDGG 60
DB 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTTHDNIITDGG 60
QY 61 IEIGDAPFOQTQIRAQAEIADEADVILFMVNVREGLTQSDENVAQQLYKSKRPVLA VNK 120
DB 61 IEIGDAPFOQTQIRAQAEIADEADVILFMVNVREGLTQSDENVAQQLYKSKRPVLA VNK 120
QY 121 VDNMEKRTVDYDFYSLGFGEPYPISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
DB 121 VDNMEKRTVDYDFYSLGFGEPYPISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
QY 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTTDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
DB 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTTDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
QY 241 ESTEKYSVLRAKAIERSNVLVVDAEGLIEODKRVAGYAHEOGKAVVIVNKKMDTYE 300
DB 241 ESTEKYSVLRAKAIERSNVLVVDAEGLIEODKRVAGYAHEOGKAVVIVNKKMDTYE 300
QY 301 KOSKTKKKFEDEVRKFEQFLDYAQIAFVSAKERTRLTLFPYINSESENHKKRVOSSTLN 360
DB 301 KOSKTKKKFEDEVRKFEQFLDYAQIAFVSAKERTRLTLFPYINSESENHKKRVOSSTLN 360
QY 361 EYVTDIAISMNPPTDGRRLNVFYATQVAIEPPTFVFNVDVLMHFSYKRYLENQIRAA 420
DB 361 EYVTDIAISMNPPTDGRRLNVFYATQVAIEPPTFVFNVDVLMHFSYKRYLENQIRAA 420
QY 421 FGEFGTPIHIIARRK 435
DB 421 FGEFGTPIHIIARRK 435
RESULT 5
ABP38338 standard; Protein; 443 AA.
XX
AC ABP38338.

XX 24-JUL-2002 (first entry)
XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3183.
XX DE
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KM antibacterial; gene therapy.
XX OS
XX Staphylococcus epidermidis.
XX PN US6380370-B1.
XX PD 30-APR-2002.
XX PF 13-AUG-1998; 98US-0134001.
XX PR 14-AUG-1997; 97US-055779P.
XX PR 08-NOV-1997; 97US-064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm LA, Bush D;
XX WPI: 2002-381255/41.
XX N-PSDB; ABN90883.
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 3183; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences
XX can also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life
XX cycle or inhibit S. epidermidis infection.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX USPTO web site.
SQ Sequence 443 AA:
Query Match 94.6%; Score 2094; DB 23; Length 443;
Best Local Similarity 93.1%; Pred. No. 1.1e-173;
Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;
QY 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTTHDNIITDGG 60
DB 8 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTTHDNIITDGG 67
QY 61 IEIGDAPFOQTQIRAQAEIADEADVILFMVNVREGLTQSDENVAQQLYKSKRPVLA VNK 120
DB 61 IEIGDAPFOQTQIRAQAEIADEADVILFMVNVREGLTQSDENVAQQLYKSKRPVLA VNK 127
QY 121 VDNMEKRTVDYDFYSLGFGEPYPISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
DB 128 VDNLEMRNDIYDFYSLGFGEPYPISGSHGLGDLDDAVVENNKSESDPYDEDTIRLSI 187
QY 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTTDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
DB 188 IGRPNVGSLLVAIIGEDRVIVSNVAGTTTDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 247
QY 241 ESTEKYSVLRAKAIERSNVLVVDAEGLIEODKRVAGYAHEOGKAVVIVNKKMDTYE 300
DB 248 ESTEKYSVLRAKAIERSNVLVVDAEGLIEODKRVAGYAHEOGKAVVIVNKKMDTYE 307
QY 301 KOSKTKKKFEDEVRKFEQFLDYAQIAFVSAKERTRLTLFPYINSESENHKKRVOSSTLN 360
DB 308 KOSKTKKKFTDDVDRNFEQFLDYAQIAFVSAKERTRLTLFPYINSESENHKKRVOSSTLN 367

QY 361 EVYTDALISMNPTPDGKRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
|||||
DB 368 EVYTDALISMNPTPDGKRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 427
QY 421 FGFECTPIHIITARRK 436
|||||
DB 428 FGFECTPIHIITARRK 443

RESULT 6

ABBA49212 ID ABB49212 standard; Protein: 436 AA.

AC ABB49212;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #1916.

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

PN WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Eslihi H, Dehoux P,
Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P,
Danjels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA;
Domiguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;
Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
Rose M, Voss H;

PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

PI Rose M, Voss H;

DR WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides

PS Claim 6; SEQ ID NO 1917, 192pp; French.

XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded from the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions and related organisms.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 436 AA;

Query Match

78.2%; Score 1731; DB 23; Length 436;

Best Local Similarity 73.6%; Pred. No. 4.4e-142;
Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIIGERVSYEDTGVYRDRIRYSGGEMITDFNITIDGG 60
|||:|||||
DB 1 MAKPVVAIVGRPNVSKSTIFNRIIGERVSYEDTGVYRDRIRYSGGEMITDFNITIDGG 60
QY 61 IEIGDAPFOTQIRAOAEIAIDEDADVITFMVNVREGLTQSDENVAGIILKSKKPPVLAANK 120
|||:|||||
DB 61 IDLSDEPPELQIRAOAEIAIDEDADVITFMVNVREGLTQSDENVAGIILKSKKPPVLAANK 120
QY 121 VDNMEMRTDYYDYSISLGEPEYPIGSHGLGLDLDLVAVSHFGEEEDPPYEDTIRLSI 180
|||:|||||
DB 121 VDNMEMRTDYYDYSISLGEPEYPIGSHGLGLDLDLVAVSHFGEEEDPPYEDTIRLSI 180
QY 181 IGRPNVGSLSLVNAILGEDRVYVNAVGTTRDAIDTEYSYDGDYVLIIDTAGMRKKGVY 240
|||:|||||
DB 181 IGRPNVGSLSLVNAILGEDRVYVNAVGTTRDAIDTEYSYDGDYVLIIDTAGMRKKGVY 240
QY 241 ESTEYSVLRLAKAIERSNVVLYVIDAEGITIEDKRVAGVAHEGKAVYIVNKMDFYE 300
|||||
DB 241 ESTEYSVLRLAKAIERSNVVLYVIDAEGITIEDKRVAGVAHEGKAVYIVNKMDFYE 300
QY 301 KDSKTKMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYINASENHKKRVOSTLN 360
|||:|||||
DB 301 KDSKTKMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYINASENHKKRVOSTLN 360
QY 361 EVYTDALISMNPTPDGKRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
|||:|||||
DB 361 EVYTDALISMNPTPDGKRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
QY 421 FGFECTPIHIITARRK 435
|||||
DB 421 FGFECTPIHIITARRK 435

RESULT 7

ABP27827 ID ABP27827 standard; Protein: 436 AA.

AC ABP27827;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4830.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

PN WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
Tettelin H;

PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

DR WPI; 2002-352536/38.

DR N-PSDB; ABN68458.

PT New Streptococcus protein for the treatment or prevention of infection
or disease caused by streptococcus bacteria, such as meningitis, and

RESU:7 9
 ABP27826
 ID ABP27826 standard; Protein; 436 AA.
 XX
 AC ABP27826;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 4828.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tetzelin H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB; ABN68457.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 Claim 1; Page 3645; 4525pp; English.
 XX
 PS The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 Sequence 436 AA;
 XX

Query Match	68.8%	Score 1523;	DB 23;	Length 436;
Best Local Similarly	66.3%	Pred. No. 5.8e-124;		
Matches 289; Conservative	72;	Mismatches 73;	Indels 2;	Gaps 2

QY 1 MTKIVAAVGRPNNGSKSTIFENRIGERSIYEDIPGVTRDRIRYSSGEWLTHTDFTNIDTGG 60
 Dd 1 MPELTVAVGRPNNGSKSTLIFENRIGERSIYEDVEGVTRDRIRYTTGEWLNRRKFSLIDTGG 60
 QY 61 IEIDAPRQPTIRKQAEIAIDEDAVYIIPMVNRRELITQSDSEWVAOILYKSKPVLAVNKK 120

Db	61	IDDVADAPMEQKKHQADIAIMTEADYIVFVVSCKEGSTDADEYVSRLTYTKNKVILAVNK	120
Qy	121	VDNMERMTRDYVDFYSLGEGEPYPISGSHGLGLGDLDAVVSHPGEEEDPYDEDTRLS	180
Db	121	VDNPEMRNDIYDFYSLGLDDPYPLSSVHGIGTGDIIDAIVENLPVEEENE-NPDIIRFSL	179
Qy	181	IGRPVNGSSLVNALGDEDRVIVSNVAGTTRRAIDPEY-SYGDQDVLLDTAGMRKKGV	239
Db	180	IGRPVNGSSLLNALGDEDRVIVASPACTTRRAIDNPFVDSOGQETMTIDTACMRKSGV	239
Qy	240	YESTEKYSVLRALKAIERSNVVLYVIDAFOGIIEDQKRVAGYAHNEGKAVIYVNNMDTV	299
Db	240	YENTKYSVMSMRKAIIDSDVLYLWVIMNAEEGRIREDKRLAGFAHETGKGIIVYNNMDTI	299
Qy	300	EKDSKTMKFEDEVEKKEFQFLDYAOIAFVSAKERTPLRTLPFYINEASENNKKRVOSTL	359
Db	300	EKDNTVTSQWEMADINDNQFLSYAPIIFVSAETKQRLHKLPDMKIRKISSQNNRIPSAVL	359
Qy	360	NEVYTDIAISNFTPPDKGRRLNVFATQVAIEPPTFFVYVNDVELMHFSYKRYLEQIRA	419
Db	360	NDVIDAIAINFTPPDKGRKRLKIFATQVAIVPPTFFVEVNEBELMHFSYRLLENQIRE	419
Qy	420	AFGEFGTPHIIARRK 435	
Db	420	AFVEEGTPINLIARRK 435	
RESULT 10			
AAB04108			
ID	AAB04108	standard: Protein; 436 AA.	
XX			
AC	AAB04108;		
XX			
DT	11-APR-2001 (first entry)		
XX			
DE	YphC protein of Streptococcus pneumoniae (GTP binding protein).		
XX			
KW	yphC: GTP binding protein; antibody; treatment; infection;		
KW	bacteraemia; otitis media; conjunctivitis; pneumoniae;		
KW	meningitis; sinusitis; pleural empyema; endocarditis.		
OS	Streptococcus pneumoniae.		
XX			
PN	WO200068427-A1.		
PD	16-NOV-2000.		
XX			
PE	02-MAY-2000; 2000WO-US11894.		
XX			
PR	07-MAY-1999; 99US-0307003.		
XX			
PA	(SMIK) SMITHKLINE BEECHAM CORP.		
PA	(SMIK) SMITHKLINE BEECHAM PLC.		
XX			
PI	Zalacain M, Biswas S, Burnham MKR, Warren PV, Ingraham KA;		
PI	Chalker AF, So CY, Holmes DJ, Warren RL, Van Horn S, Trainl CM;		
XX			
DR	WPI: 2000-687653/67.		
XX			
DR	N-PSDB: AAA54516.		
XX			
PT	Streptococcus pneumoniae yphC protein and DNA sequence, useful for		
PT	treating infections, meningitis, and bacteraemia		
XX			
PS	Claim 1; Page 37-38; 39pp: English.		
XX			
CC	A new nucleotide is described which encodes a 436 residue		
CC	Streptococcus pneumoniae yphC (GTP binding protein). The DNA		
CC	sequence can be used to transform a host cell to produce		
CC	the protein and the products can be used to treat bacterial		
CC	infections (especially Streptococcus pneumoniae infections, and		
CC	Helicobacter pylori infections), otitis media, conjunctivitis,		
CC	pneumonia, bacteraemia, meningitis, sinusitis, pleural empyema, and		
CC	endocarditis. The encoded protein and its peptide fragments can be		

Streptococcus pneumoniae yphc protein and DNA sequence, useful for treating infections, meningitis, and bacteraemia

A new nucleotide is described which encodes a36 residue Streptococcus pneumoniae ypic (GTP binding protein). The DNA sequence can be used to transform a host cell to produce the protein and the products can be used to treat bacterial infections (especially Streptococcus pneumoniae infections, and Helicobacter pylori infections), otitis media, conjunctivitis, pneumoniae, bacteraemia, meningitis, sinusitis, pleural empyema, endocarditis. The encoded protein and its peptide fragments can be

KW	ear infection; antiinflammatory; antibacterial; immunostimulant;
KW	auditory; respiratory; gene therapy; vaccine.
XX	
OS	Streptococcus pneumoniae type 4 strain.
XX	
PN	WO200277021-A2.
XX	
PD	03-OCT-2002.
XX	
PF	27-MAR-2002; 2002WO-IB02163.
XX	
PR	27-MAR-2001; 2001GB-0007658.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	(GENO-) INST GENOMIC RES.
XX	
DR	Maignani V, Tettelin H, Fraser C;
DR	MP1: 2003-040579/03.
DR	N-PSDB: ABX07474.
XX	
PT	New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT	useful as medicaments for treating or preventing a disease or infection
PT	due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
PT	or ear infection -
XX	
CS	Claim 1; SEQ ID No 3524; 56pp; English.

	Query Match	Match Similarity	Best Local	Conservative	Mismatches	Score	DB	length
	290	68.8%	65.2%	73	69	1522	24	436
QY	1	MTKPLVIAVGRPNCKSTLFNRI	RIYERISYVDTCGVDRDRISYSGEWLTHFN	ILIDTG	60			
DB	1	MALPITALVGRPNCKSTLFNRI	RIYERISYVDTCGVDRDRISYSGEWLTHFN	ILIDTG	60			
QY	61	IEIGDAPQOTQIRAOAETAI	DAVDYITFMVNVBSGLTSD	DEKVAQILYKSKRPVLA	120			
DB	61	IDDDVAPMEQKIKHOAEI	AMEADYIVVSKSEIITAD	EVAKRLKTHRPVLA	120			

```

0Y 121 VDNMRRDHYVEYFSGFEPEYPISSSHGLGJLDLADSHQGE--EEDPDPTDRTIRL 178
    ||| ||| : |||| : ||| ||| ||| ||| ||| : ||| ||| : ||| ||| : |||
Db 121 VDNPEMRNDIDYFALGLGEEPLPISSVHEIGTGDDVALVENLPNEYEBBNP--DYIKF 177

0Y 179 SLIGRPNVGKSSLLVAALIGEDRVIVSNVAGTTTDAIDTETS--YDGODYVLIDPAGRRKG 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 SLIGRPNVGKSSLLVAALIGEDRVIVSNVAGTTTDAIDTFTDTDGGEFIMIDPAGRRKS 237

0Y 238 KYESTERYSVLRALKAIERSNVVLVIDAEOGIIEDOKRVAGYAHEOGKAVYIVVKKWD 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 KYENTEKYSVWRAMRAIDRSDVVLVIVNAEGIREYDKRIAGEFAHEAGKMIIVVKKWD 297

0Y 298 TVERKSKTMKKEDEYKRKEFOFLDYAQLAFVSAKETRLTLPEYINSESENKKRVQSS 357
    ||| ||| ||| : ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 298 TLEKNHMTKMKNEEDIREQFOYLPAPPIIFVSALTRQHLKHPLEMKIJOISENTRIPSA 357

0Y 358 TLNEVVTGTAISNPPIDOKGRRLNFEATOVALIEPTFVAVVDVDELHMSYKRYENOI 417
    ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 VLNDYIMAIALINPPIDOKGRKLKIFYATOVATKPTPEVIFVNEELMHPSYTLRENOI 417

0Y 418 RAAFEGEGTPIHIIRKKR 435
    || || ||||| |||||
Db 418 RKAPEEGTPIHILIRKKR 435

```

RESULT13	
ABB54068	
ID	ABB54068 standard; Protein; 436 AA
XX	
AC	ABB54068;

Query Match 68.7%; Score 1520; DB 23; Length 436;
Best Local Similarity 66.3%; Pred. No. 1,1e-123;
Matches 289; Conservative 69; Mismatches 76; Indels 2; Gaps 2;

QY 1 MTKPIVALVGRPNVSKSTIFNRIVGERVSIYEDTFCGYRDRIRYSSGEMLTHTDFNIIDGG 60
DB 1 MSLPIVALVGRPNVSKSTIFNRIVGERVSIYEDTFCGYRDRIRYATGEMLTHTDFNIIDGG 60
QY 61 IEIGDAPFOQIRAOAEIAIDADVIFMVNVRGLTQSDENVAQILYKSKRPVLAVN 120
DB 61 IELSDPEPTEIRAQAEIATMEADYIAVVDGEGTGTDADEAVANILYRTDKPVLIVNKK 120
QY 121 VDNMERTDVFDFSLGCEPPISSGSLGLDLDLDAVSHFGEEDDPDEDTRISI 180
DB 121 VDNPRRMEIFDFYSLGLDPPYPSAVHIGIGDVLAIYONLPNEIEEE-NEENIKFSL 179
QY 181 IGRPNVSKSLVNAITLGEDRVIVSNAGTTRDAIDTEX-SYDGODYVIDTPAGMKKKCV 239
DB 180 IGRPNVSKSLVNAITLGEDRVIVSNAGTTRDAIDTEX-SYDGODYVIDTPAGMKKKCV 239
QY 240 YESTEKYSVLRAKAIERSNVLVVIADEOGIIEODKRVAGYAHOGKRAVIVVNMKMDTV 299
DB 240 YENFEKYSVNRAMRAIDRSIDITLVYINAEIGIREYDMHAGFAHAGKGLILVNMKMDTV 299
QY 300 EKDSITMKKFEDEVKKEFOFLDYAOIAFVSAKERTRLTLPYINEASENHKKRVQSTL 359
DB 300 EKDNVTMNFELERTKERFLDYAPIYVSAKTGRLKLPDIKEIHHQNLRISSSVL 359
QY 360 NEVVTDAISMNPPTDKGRNLNVEFATQVAIEPPEFVVVNDVELMHFSYKRYLENOIRA 419
DB 360 NDVIMDAVINPTPTDKGRNLKIFATQVAIKPPEFVVVNDVELMHFSYKRYLENOIRK 419
QY 420 AFEGEGTPIHIIARR 435
DB 420 AFEGEGTPIHIIARR 435

RESULT 14
AAG99939 ID AAG99939 standard; Protein: 448 AA.
XX AC AAG99939;
XX DT 27-SEP-2001 (first entry)
XX DE ERA binding domain polypeptide SEQ ID NO 381.
XX KM ERA binding domain; Escherichia coli; GRPase; antimicrobial;
KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
XX peptide therapy.
XX OS Mycobacterium leprae.
XX PN WO200153458-A2.
XX PD 26-JUL-2001.
XX PF 17-JAN-2001; 2001WO-US01786.
XX PR 18-JAN-2000; 2000US-0176870.
XX RA (SMIK) SMITHKLINE BEECHAM CORP.
XX RA (SMIK) SMITHKLINE BEECHAM PLC.
XX PI Lupas AN, Pearce KH;
XX DR WPI: 2001-476108/51.
XX PT New ERA binding domain polypeptides and polynucleotides encoding them,
XX PT useful as research reagents and materials for discovery of treatments
XX PS and diagnostics for diseases, or for genetic immunisation -
PS Claim 1; Page 52-53; 279pp; English.

XX The present invention relates to ERA binding domain polypeptides
CC (AAG99939-AAG99989 and AAM00010-AAM00021). The era gene in Escherichia
CC coli codes for an essential GRPase protein able to autophosphorylate at
CC serine and/or threonine residues. The protein has potential antimicrobial
CC and antibacterial activity and is useful in screening for antagonists,
CC agonists and for compounds with antibiotic activity. The proteins are
CC also useful in determining their role in pathogenesis of infection,
CC dysfunction and disease and could be used as part of a vaccine and/or
CC peptide therapy.
SQ Sequence 448 AA:

Query Match 44.9%; Score 933; DB 22; Length 448;
Best Local Similarity 45.6%; Pred. No. 8.7e-78;
Matches 199; Conservative 93; Mismatches 136; Indels 8; Gaps 6;

QY 1 MTKPIVALVGRPNVSKSTIFNRIVGERVSIYEDTFCGYRDRIRYSSGEMLTHTDFNIIDGG 60
DB 1 MSLPIVALVGRPNVSKSTIFNRIVGERVSIYEDTFCGYRDRIRYATGEMLTHTDFNIIDGG 58
QY 61 IEIG-DAPFOQIRAOAEIAIDADVIFMVNVRGLTQSDENVAQILYKSKRPVLAVN 119
DB 59 LVFNDDSEFLPEIRQANLA--EAKAALFVVDGQGPPTASDEIQAOLROOSVPIYIAVN 116
QY 120 KVDNMEM-RTDVFDFSLGCEPPISSGSLGLDLDLDAVSHFGEEDDPDEDTRIL 178
DB 117 KCESPQGAIQAAEFMHGLGEPYMSAIIHSGTGDLDALEYLELPAPQEP-EEDEIKV 175
QY 179 SIIGRPNVKSSLVNAITLGEDRVIVSNAGTTRDAIDTEXSYDGODYVLTITAGARRKKG 238
DB 176 AIVGRPNVSKSLVNAITLGEDRVIVSNAGTTRDAIDTEXSYDGODYVLTITAGARRKN 235
QY 239 YESTEKYSVLRAKAIERSNVLVVIADEOGIIEODKRVAGYAHOGKRAVIVVNMKMDT 298
DB 236 VDYGAEFFGIRNRAFAIRADVYLFVLDVLDGVTBQDLKAGRIEDRAVAVVINKMDA 295
QY 299 VEKDSITMKKFEDEVKKEFOFLDYAOIAFVSAKERTRLTLPYINEASENHKKRVQST 358
DB 296 VEKDSYTIYERHQLMARLYFMDWAEIMFVSAQGLRVOKILDCVDIAAOEHRRRVYTA 355
QY 359 LNEVVTDAISMNPPTDK-GRRLNVEFATQVAIEPPEFVVVNDVELMHFSYKRYLENOI 417
DB 356 INEVLAEAVSWHSPPTTRQKQKGIYGTGVSTOPPALALFVNDPRNDVRYIEKQF 415
QY 418 RAAFEGEGTPIHIIARR 433
DB 416 RKQLGFGSPTRILFWR 431

RESULT 15
AAG99941 ID AAG99941 standard; Protein: 448 AA.
XX AC AAG99941;
XX DT 27-SEP-2001 (first entry)
XX DE ERA binding domain polypeptide SEQ ID NO 383.
XX KM ERA binding domain; Escherichia coli; GRPase; antimicrobial;
KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
XX peptide therapy.
XX OS Treponema pallidum.
XX PN WO200153458-A2.
XX PD 26-JUL-2001.
XX PF 17-JAN-2001; 2001WO-US01786.
XX PR 18-JAN-2000; 2000US-0176870.

XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Lupas AN, Pearce KH;
XX WPI: 2001-476108/51.
DR
XX
XX
PT New ERA binding domain polypeptides and polynucleotides encoding them,
PT useful as research reagents and materials for discovery of treatments
PT and diagnostics for diseases, or for genetic immunisation -
XX
PS Claim 1; Page 53; 279pp: English.
XX
XX The present invention relates to ERA binding domain polypeptides
CC (AAG99559-AAG99989 and AAM00010-AAM00021). The era gene in Escherichia
CC coli codes for an essential GTPase protein able to autophosphorylate at
CC serine and/or threonine residues. The protein has potential antimicrobial
CC and antibacterial activity and is useful in screening for antagonists,
CC agonists and for compounds with antibiotic activity. The proteins are
CC also useful in determining their role in pathogenesis of infection.
CC dysfunction and disease and could be used as part of a vaccine and/or
CC peptide therapy.
XX
XX
SQ Sequence 448 AA;

Query Match: 42.3%; Score 935.5; DB 22; Length 448;
Best Local Similarity 43.2%; Pred. No. 8.8e-73;
Matches 189; Conservative 86; Mismatches 134; Indels 9; Gaps 4;

QY 6 VAIIGRPNVGKSTIFNRIYGERVSIVEDTPGTRDIYSSGEMLTDFNIIDTGGIEIGD 65
Db 4 VAIIGRTNVGKSTIFNRIQKPAIVSDPTPTTRDRIFGIGEMLRKRIAFIDTGLLAKQ 63

QY 66 APEGTOIRAOAEIIDEADVITFMVNVREGLTQSDENYAQILYKSK-KPVVLAIVKVDNM 124
Db 64 TPLOQLALQVQAKIIFLVSLOEQLSDDFYVAKVLRKKNDKPVILVYVNAKENF 123

QY 125 EMRT---DYDYFYSLGFGPEPYISGSHGLGDLDAVYSH---FGEEEDPYDEDTRL 178
Db 124 NPKTAETLNDYISLGFGRPVYISAHIGIGIDMLDLVKONOLLPNEND--DIAKITF 181

QY 179 SIIGRPNVGKSSILVNAIIGEDRVIVSNVAGTTTDAIDTEYSYGODVYLIDTAGMKRKGK 238
Db 182 CVIGKPNVGKSSILNQVQKRVLYVNSGTTTDAIDVPLKVNCEKFLIDTAGIKRKGK 241

QY 239 VYESTEKYSVLRALKATERSNVVLVIDAEGITIEQDKRVAGYAHGCKAVYIVYKWDT 298
Db 242 INMGIEIYASYIKTILAIRSNVILLMWDGSKPISEODEVIGLAQALIPVILVKNKWL 301

QY 299 VEKDSKTMKKRFEDEVKKEFOFLDYAQIAFVSAKERTRLTPYINEASENHKKRVOSST 358
Db 302 VLKNNNTNAYKKMLKLFKHLDRAPVLFISVLKNQRLNTLFEOLKITQSOLETVATPL 361

QY 359 LNEVVTDAISMNPPTDKGRRLNLYFAIQVAIEPPTFVFNVDVLMHFSYKRYLENQIR 418
Db 362 LNDVYIQAAQLNQPRLFKGRRLQITTVANQTSQIRPHFVLCNDPKYILHFSYARLENKIR 421

QY 419 AAFGEETPIHIIARKRN 436
Db 422 ENFGFNSVPIISLYFKSKN 439

Search completed: September 10, 2003, 00:32:44
Job time : 88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:51 ; Search time 43 Seconds

(without alignments)
975.106 Million cell updates/sec

Title: US-09-815-242-12600

Perfect score: 2213

Sequence: 1 MTKPIVAIVGRPNVGKSTIF.....IRAFGEFGPIHIARKRN 436

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2213	100.0	436	2 C89926	hypothetical prote
2	1731	78.2	436	2 A11316	hypothetical prote
3	1724	77.9	436	2 A11688	hypothetical prote
4	1699	76.8	436	2 A69936	probable GTP bindi
5	1658	74.9	437	2 F83854	hypothetical prote
6	1522	68.8	436	2 B95199	phosphoglycerate d
7	1522	68.8	436	2 H98065	phosphoglycerate d
8	1520	68.7	436	2 C86719	GTP-binding protei
9	1218.5	55.1	438	2 B97111	probable GTPase (I
10	1037	46.9	433	2 A81867	GTP binding protei
11	1021	46.1	442	2 S75645	probable GTP bindi
12	995	45.0	442	2 B82899	conserved GTP bindi
13	984.5	44.5	435	2 H90522	gtp-binding protei
14	935.5	42.3	448	2 D64236	hypothetical prote
15	923	41.7	503	2 E91050	probable GTP-bind
16	923	41.7	503	2 A85895	probable GTP-bind
17	923	41.7	503	2 F65027	probable GTPase/GT
18	910	41.1	490	2 AF0821	probable GTP-bind
19	906.5	41.0	490	2 B82284	GTP-binding protei
20	896.5	40.5	495	2 AC0350	probable GTP-bind
21	864.5	39.1	483	2 A70465	essential GTP bindi
22	860.5	38.9	485	2 G81149	probable GTP-bind
23	858.5	38.8	485	2 E81871	probable GTP-bind
24	852	38.5	504	2 F64143	probable GTPase/GT
25	842.5	38.1	439	2 B72253	conserved GTP-bind
26	840.5	38.0	493	2 B83171	conserved GTP-bind
27	838	37.9	493	2 D87454	GTP-binding protei
28	806.5	36.4	463	2 H70504	probable GTP-bind
29	797.5	36.0	483	2 AH3445	GTP-binding protei

30	789	35.7	476	2 A97636	probable GTP-bind
31	789	35.7	476	2 AC2859	GTP-binding protei
32	784	35.4	438	2 F75290	conserved GTP-bind
33	782.5	35.4	461	2 S72953	probable GTP-bind
34	782.5	35.4	462	2 F87080	probable GTP-bind
35	779	35.2	465	2 D82804	GTP-binding protei
36	778.5	35.2	476	2 A95982	probable GTP-bind
37	752	34.0	453	2 C85000	hypothetical GTP-b
38	746.5	33.7	447	2 F97828	conserved GTP-bind
39	737.5	33.3	447	2 H71673	hypothetical prote
40	734.5	33.2	460	2 H81381	probable GTP-bind
41	719.5	32.5	490	2 A81744	GTP-binding protei
42	710.5	32.1	490	2 H71480	probable GTPase/GT
43	706	31.9	458	2 B64624	GTP-binding protei
44	701	31.7	462	2 E71891	probable GTP bindi
45	685.5	31.0	487	2 B86596	GTPase/GTP-binding

ALIGNMENTS

RESULT 1	C89926	hypothetical protein SA1307 [imported] - Staphylococcus aureus (strain N315)
C:Species:	Staphylococcus aureus	
C:Date:	10-May-2001	#sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession:	C89926	
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O		
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.		
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.		
Lancet 357, 1225-1240, 2001		
A:Title:	Whole genome sequencing of methicillin-resistant Staphylococcus aureus.	
A:Reference number:	A89758; MUID:21311952; PMID:11418146	
A:Accession:	C89926	
A:Status:	Preliminary	
A:Molecule type:	DNA	
A:Cross-references:	GB:BA000018; PID:g13701274; PIDN:BAH42568.1; GSPDB:GN00149	
A:Experimental source:	strain N315	
C:Gene:	SA1307	
C:Superfamily:	Mycobacterium leprae probable GTP-binding protein; translation elongat	
Query Match	100.0%; Score 2213; DB 2; Length 436;	
Best Local Similarity	100.0%; Pred. No. 1e-128;	
Matches	436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 MTKPIVAIVGRPNVGKSTIFNRIYGERYSIVEDPFGVTRDRIRYSGEWLTHDFNIIDTGG 60	
DB	1 MTKPIVAIVGRPNVGKSTIFNRIYGERYSIVEDPFGVTRDRIRYSGEWLTHDFNIIDTGG 60	
OY	61 IEIGAPQPTQIRAOAEIAIDEDAVIIFMVNVRBSLTQSDENVAOILYKSKPPVLAANK 120	
DB	61 IEIGAPQPTQIRAOAEIAIDEDAVIIFMVNVRBSLTQSDENVAOILYKSKPPVLAANK 120	
OY	121 VDNMEMRVDYDFYSLGEGEPYISGSHGLGDLDAVVSFGGEEDPYEDTIRLSI 180	
DB	121 VDNMEMRVDYDFYSLGEGEPYISGSHGLGDLDAVVSFGGEEDPYEDTIRLSI 180	
OY	122 VDNMEMRVDYDFYSLGEGEPYISGSHGLGDLDAVVSFGGEEDPYEDTIRLSI 180	
DB	122 VDNMEMRVDYDFYSLGEGEPYISGSHGLGDLDAVVSFGGEEDPYEDTIRLSI 180	
OY	181 IGRPNVKSLSLVALNLGDERVIVSNVAGTTRDAIDTEXSYDGDVLLDTAGMRKKGVY 240	
DB	181 IGRPNVKSLSLVALNLGDERVIVSNVAGTTRDAIDTEXSYDGDVLLDTAGMRKKGVY 240	
OY	182 IGRPNVKSLSLVALNLGDERVIVSNVAGTTRDAIDTEXSYDGDVLLDTAGMRKKGVY 240	
DB	182 IGRPNVKSLSLVALNLGDERVIVSNVAGTTRDAIDTEXSYDGDVLLDTAGMRKKGVY 240	
OY	241 ESTEYYSVLRALKAIERSNVVLVDAEAGIIEQDKRVAGYAHGOKAVIYVNWDTVE 300	
DB	241 ESTEYYSVLRALKAIERSNVVLVDAEAGIIEQDKRVAGYAHGOKAVIYVNWDTVE 300	
OY	301 KOSTMKKFEDEVREKPEFLDVAQIAFASAKERTLRLLPYINSESHKKRVOSTLN 360	
DB	301 KOSTMKKFEDEVREKPEFLDVAQIAFASAKERTLRLLPYINSESHKKRVOSTLN 360	
OY	361 EVVTAISNNPPTKGRRLNVFATQVAIEBPFTVVVNDVLMHFSKRYLENOIRAA 420	
DB	361 EVVTAISNNPPTKGRRLNVFATQVAIEBPFTVVVNDVLMHFSKRYLENOIRAA 420	

Db 361 EVVTDALISNNPTPTDGRRLNFFATQVAIEPPTFVFNVDVLMHFSTKRYLENDQIRAA 420
 QY 421 FGEFGTPIHIARKRN 436
 Db 421 FGEFGTPIHIARKRN 436

RESULT 2

hypotheetical protein lmo1937 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: A11316
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefl, U.; Kuhn, M.; Kunst, F.; Kurapkát, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Thieriez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11316
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00015.1; PID:g16411390; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1937
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

Query Match 78.2%; Score 1731; DB 2; Length 436;
 Best Local Similarity 73.6%; Pred. No. 4.3e-99;
 Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

QY 1 MKKPIVAIVGRPNVKGSTIFNRIVGRVSVIEDTPGVTRDRIYSSEGLTHDNIITDGG 60
 Db 1 MAKPVVAIVGRPNVKGSTIFNRIVGRVSVIEDTPGVTRDRIYSSEGLTHDNIITDGG 60
 QY 61 IEIGDAPFQOTIRAOAEIAIDADVIFVNVNREGTQSDENVAQILYKSKRPVLAIVK 120
 Db 61 IDLSDEPFLEQIRAOAEIAIDADVIFITNGREGVTADAEQAKILYKSKRPVLAIVK 120
 QY 121 VNMENKRTVDYFYSIGFGEPPYPISSHGLGGLLDAAVSHFGEEDPVEDTIRLSI 180
 Db 121 VNPEKRDQIDYFYSIGFGEPPYPISSHGLGGLMDIAVRAHPKKEEEDPVEDTIRKFSL 180
 QY 181 IGRPNVKGSSILNALGEDRVIVSNVAGTTRADIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Db 181 IGRPNVKGSSILNALGEDRVIVSDIAGTTRADIDTTFDQDQVYIMDTAGMRKRGKY 240
 QY 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHOGKAVVIVNKKDVE 300
 Db 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHOGKAVVIVNKKDVE 300
 QY 301 KDSKTKKREDEVRKPEFQFLDYAOIAFSAKERTRLTFPYINSESHKRRVOSTLN 360
 Db 301 KDEKTIINWTEIDIRDFQFLSTAPLVFSAKTKRNLNPLINOVDHSHSRVSSMLN 360
 QY 361 EVVTDALISNNPTPTDGRRLNFFATQVAIEPPTFVFNVDVLMHFSTKRYLENDQIRAA 420
 Db 361 DVISDAVAMNPSPMOKGRKRLKIFTTQYAVAKPPTVFNVDVLMHFSTKRYLENDQIRAA 420
 QY 421 FGEFGTPIHIARKRN 435
 Db 421 FFEFGTPIHIARKRN 435

RESULT 3
 A11688
 hypotheetical protein homolog lln2051 [imported] - Listeria innocua (strain Clifp11262)
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: A11688
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefl, U.; Kuhn, M.; Kunst, F.; Kurapkát, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Thieriez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11688
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97281.1; PID:g16414552; GSPDB:GN00178
 A:Experimental source: strain Clifp11262
 C:Genetics:
 A:Gene: lln2051
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat

Query Match 77.9%; Score 1724; DB 2; Length 436;
 Best Local Similarity 73.1%; Pred. No. 1.2e-98;
 Matches 318; Conservative 67; Mismatches 50; Indels 0; Gaps 0;

QY 1 MKKPIVAIVGRPNVKGSTIFNRIVGRVSVIEDTPGVTRDRIYSSEGLTHDNIITDGG 60
 Db 1 MAKPVVAIVGRPNVKGSTIFNRIVGRVSVIEDTPGVTRDRIYSSEGLTHDNIITDGG 60
 QY 61 IEIGDAPFQOTIRAOAEIAIDADVIFVNVNREGTQSDENVAQILYKSKRPVLAIVK 120
 Db 61 IDLSDEPFLEQIRAOAEIAIDADVIFITNGREGVTADAEQAKILYKSKRPVLAIVK 120
 QY 121 VNMENKRTVDYFYSIGFGEPPYPISSHGLGGLLDAAVSHFGEEDPVEDTIRLSI 180
 Db 121 VNPEKRDQIDYFYSIGFGEPPYPISSHGLGGLMDIAVRAHPKKEEEDPVEDTIRKFSL 180
 QY 181 IGRPNVKGSSILNALGEDRVIVSNVAGTTRADIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Db 181 IGRPNVKGSSILNALGEDRVIVSDIAGTTRADIDTTFDQDQVYIMDTAGMRKRGKY 240
 QY 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHOGKAVVIVNKKDVE 300
 Db 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHOGKAVVIVNKKDVE 300
 QY 301 KDSKTKKREDEVRKPEFQFLDYAOIAFSAKERTRLTFPYINSESHKRRVOSTLN 360
 Db 301 KDEKTIINWTEIDIRDFQFLSTAPLVFSAKTKRNLNPLINOVDHSHSRVSSMLN 360
 QY 361 EVVTDALISNNPTPTDGRRLNFFATQVAIEPPTFVFNVDVLMHFSTKRYLENDQIRAA 420
 Db 361 DVISDAVAMNPSPMOKGRKRLKIFTTQYAVAKPPTVFNVDVLMHFSTKRYLENDQIRAA 420
 QY 421 FGEFGTPIHIARKRN 435
 Db 421 FFEFGTPIHIARKRN 435

RESULT 4
 A69936
 probable GTP binding protein yphc - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
 C:Accession: A69936
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Entlian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizitz, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koelter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scani A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

data no good

akeuchi, M.; Takahashi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosto, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zunshtein, E.; Yoshikawa, K.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69936
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-436 <KUN>
A:Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14200.1; PID:g2634702
A:Experimental source: strain 168
A:Genetics:
A:Gene: yphC
C:Superfamily: Mycobacterium *leprae* probable GRP-binding protein; translation elongation
C:Keywords: duplication; GRP binding; nucleotide binding; P-loop
F:4-122/Domain: translation elongation factor Tu homology <ET1>
F:10-17/Region: nucleotide-binding motif A (P-loop)
F:119-122/Region: GRP-binding NKXD motif
F:145-147/Region: GRP-binding SAK/L motif
F:175-287/Domain: translation elongation factor Tu homology <ET2>
F:182-189/Region: nucleotide-binding motif A (P-loop)
F:294-297/Region: GRP-binding NKXD motif
F:329-331/Region: GRP-binding SAK/L motif

Query Match	76.8%	Score 1699	DB 2	Length 436
Best Local Similarity	73.6%	Pred. No. 4e-97		
Matches 320	Conservative 57	Mismatches 58	Indels 0	Gaps 0
QY	1	MTKPIVAIVGHPNVGSKSTIFNRKIVGERVSIYEDTPGVRTDRIVSSGEMLTNDFNIIDTGG	60	
Db	1	MGKPVVAIVGHPNVGSKSTIFNRKIVGERISYIEDPGVRTDRIVSSAEMLVNDFNLIDTGG	60	
QY	61	IEIDAPFQOQIRQAETAIIDEADVIIIPMVNREGSLQSDENMAQIIYKSKRPVYLVNKK	120	
Db	61	IDIDDEPLAIDRQAETIADEADVIIIPMVNREGSVTADEEVAKIIYRKKRPVYLVNKK	120	
QY	121	VDNNEMRTDVIYDFSLGFGEPYPIGSGHGLGLGLDLAVVNSHGEDEEDYDEDDTIRLSI	180	
Db	121	LDNTEMARNITDVFSLGFGEPYPIGSGHGLGLDLDAVAEHFNKIPETRYNEEVIGFCL	180	
QY	181	IGRPNVKSSLYVNAIIIGEDRYIVSNVAGTTRDAIDTSEYSDGQYVYLIDTAGMKRKKRY	240	
Db	181	IGRPNVGSSLYVNAIIMLEERYIVSNVAGTTTDAVDTSFYNAQGEVYLVDTAGMKRKKRY	240	
QY	241	ESTERYSVLRALKAIEKSNVYLVVIDAEQGLIEEDKKVAGYAHBOGKAVYIVVAKMPTVE	300	
Db	241	ETTEKYSVLRALKAIIDREVVAVYVLDGEGEIIEDDKRLAGYAHGAKAVYIVVAKMAYD	300	
QY	301	KDSITMKKFEDEVAKKEQFLDYAOIAVSAKERTRLTLPEYINAESENHKKRVOSSTLN	360	
Db	301	KDESTMKFEENIHDHFOFLDYAPILFMSALTKRKIRHTLMPALIIKASENSLSLRQTVNLN	360	
QY	361	EYVDAISMPNTPDKGRRLNVFVATQVAOIAPEPPVYFVNDVELMHSYKRYLENOIRAA	420	
Db	361	DVIDDAVAMNPTPHNSRLKIYATQVSVKPPSFVYVNDPELMHSTYERFLENIRIDA	420	
QY	421	FGEGSTPIHIIARRK 435		
Db	421	FGEGSTPIKIFARRK 435		

RESULT 5
F83854
hypothetical protein BH1638 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83854
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: F83854
A:Status: preliminary

A: Molecule type: DNA
A: Residues: 1-437 <STO>
A: Cross-references: GB:AP001512; GB:BA000004; MID:10174030; PIDN:BA05357.1; GSPDB:G
A: Experimental source: strain C-125
C: Genetics:
A: Gene: BH1638
C: Superfamily: Mycobacterium leprae probable GRP-binding protein; translation elongat
Query Match 74.9%; Score 1658; DB 2; Length 437;
Best Local Similarity 70.9%; Pred. No. 1.3e-94;
Matches 309; Conservative 67; Mismatches 60; Indels 0; Gaps 0;

Query Match	74.9%: Score 1658; DB 2; Length 437;
Best Local Similarity	70.9%: Pred. No. 1.3e-94;
Matches	309; Conservative 67; Mismatches 60; Indels 0; Gaps 0;
QY	1 MTKPVALVGRPNVGKSTIFNRIVGERVSVIEDTPGVTRDRIYSGSEMLTHDFNIIDTGG 60
DB	1 MSKPVIALVGRPNVGKSTIFNRIVGERVSVIEDTPGVTRDRIYSHGEMLRREFVVDITGG 60
QY	61 IEIGDAPRQOIRRAQAEIADDEADVIIFPMVNVREGGLQSDSEMAOILYKSKKRPVLAVNK 120
DB	61 IEIGDEPLVVMKRAQAEIADKADVIIFVYNNRGCVTAADQEVAKLLFRSKKRPVLGVNK 120
QY	121 VDNMMEKRTDVDFPLSGFGEPYPISGSHGLGLGLDAAVNVSHFGEEDEEDYDEDTJRLST 180
DB	121 IDHPDMQDELYEFLSGIGDPIPISGAHGLGLGLDADCAVNHPEDEGDDYDEDTJRLSL 180
QY	181 IGRPNVGKSSLVYNAVLGEDRYIVSNVAGTTRDAIDTETSYDGDQDYLIDTAGMKRKKVY 240
DB	181 IGRPNVGKSSLVYNAVLGEERYIVSNIGTTRDAIDTAFSRDQDEYVLIDTAGMKRKKVY 240
QY	241 ESTEKYSVLRLAKAIERSNVVLVYIDAEOGIIEDDKRVAGVABOGKAVYIVNKKMPTVE 300
DB	241 ESTEKYSVLRLAKAIERSDVLVYLVNEEGIIEDDKIAGVABHAGAAIIIVNKKMAYE 300
QY	301 KDSYTKMKKEDEVAKREFQFLDYAOIAFVASKERTRLTLPPYIVNEASENNKKRVQSTLN 360
DB	301 KDDPTLHNFQOKRIDEEOFLDYAPVLVESAKTKORLOHVLPAVKKVSENNLRBPTHYLN 360
QY	361 EVMVDATISMNPTPDGKGRLLNVFATGVADVAIEPFFVYVNDVLELMHFSYKRYLENOIRAA 420
DB	361 DLVMDAAVAMNPTPDHGKRLKINLYVQVAAVGPFFVYFVNDPELMHFSYARPLENRLDT 420
QY	421 FGEECTPIHIITARRKN 436
DB	421 FEFECTPIKIITARRKN 436

RESULT 6
 B95199
 phosphoglycerate dehydrogenase-related protein [imported] - Streptococcus pneumoniae
 C.Species: Streptococcus pneumoniae
 C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C.Accession: B95199
 R.Tseltelin, H.E. Nelson, K.E.; Paulsen, I.T.; Elsen, J.A.; Read, T.D.; Peterson, S.; H
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A.Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 A.Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A.Reference number: A95000; MUID:21357209; PMID:11463916
 A.Accession: B95199
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-436 <KUP>
 A.Cross-references: GB:AF005672; PIDN:AAK7587.1; PID:g14973204; GSPDB:GN00164; TIGR:
 A.Experimental source: strain TIGR4
 C.Genetics:
 C.Gene: SPI709
 C.Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat
 Query Match 68.88; Score 1522; DB 2; Length 436;
 Best Local Similarity 66.28; Pred. No. 3e-86;
 Matches 290; Conservative .73; Mismatches 69; Indels 6; Gaps 3;

Query Match	68.8%	Score 1522	DB 2	Length 436
Best Local Similarity	66.2%	Pred. No. 3e-86		
Matches 290	Conservative .73	Mismatches 69	Indels 6	Gaps 3

```

1  MALEPFAIVGRPNVNGKSTLEPNRIAGERISIVIEVEGYTRDRIVATGEMLRKNSMIDTGG 60
Qy 1  IEIGDABFQOTQIRAOAEIAIDEDADVIEIFMVNVREGLTQSDENVAQOILIKSKKEPVIAVVK 120
Db 61  IDVDVDAEFMQIKHQAEIAIMEADVIVFVVSCKEGITDADEYARLRKLYTKHKVILAVVK 120
Qy 121  VDNKEMRTDVYDFYSLGSGEPYPISSGSHGIGLDDLDAVVSHPGE--EEDPYDEDTIRU 178
Db 121  VDNEMEMRNDIYDFALDGLGEPPLISVHGIGTGDVDAIENLPNEEENP---DVIKF 177
Qy 179  SIIGRPVNGKSSLVNALILGEDRVIVSVACTTDAIDTELS-IDGQDVIIDPAGKRKKG 237
Db 178  SLIGRPVNGKSSLLNALILGEDRVIVASVACTTDAIDTHTDIDGGEFTMIDPAGKRKSG 237
Qy 238  KYVESTEKYSULRKAIEKRSNVLIYVYDMEOGITBODKRUVAGYAHOGCAVYIVVYVKMD 297
Db 238  KYVENTEKYSYMRMRARIDRSDDVYLAWYINAEBSGIREYDKRIASFAHDAGGMITIVYVKMD 297
Qy 298  TVEKDSKTMKKEFEDEVKEFEQFLDYQOIAVSAKERTRLTLEPPYINEASENHKRRVQSS 357
Db 298  TLEKDNTMTNMWEDIDIEQFOYLPYALIIIVSALITQRHLKLPDEMUKOISESNTNTIIPSA 357
Qy 358  TLNEVVTDAISMNPTDQGRRLNVFYATOVALEPTEFVYVNDVDELMHFSYKRYLENOI 417
Db 358  VLNDVINDAIAINPTPDKGRKLRIKFIYATQVATKRPTEVIFVNEEELMHFSYRPLENOI 417
Qy 418  RAAGFEGCTPIHITAKR 435
Db 418  RKAFFEGCTPIHLIARKR 435

```


305 TMIEYEKKLEKMEPFISMPIVFIISALKSQRLNKLKEVITIIQVKNNSIRELKNÜLNDLVL 364

QY 365 DAISMNPTPDGRLNVFYATQVAIEPPTFVVVFVNDVLMHFSYKRYLENOIRAAFGFE 424

Db 365 DMOQMNPDL-TFKGKKLEIKHIKTKTNDVPPEFLFVNPNIVHFSYLENIQNDIRDYDFLT 423

QY 425 GTFPHITARK 434

Db 424 GCPINLVLRK 433

RESULT 14

D64236

hypothetical protein homolog MG329 - Mycoplasma genitalium

C:Species: Mycoplasma genitalium

C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001

C:Accession: D64236

R:Rafter, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J. C.A.; Venter, J.C.

Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:7565993

A:Accession: D64236

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-448 <TIGR>

A:Cross-references: GB:U9715; GB:L43967; NID:g1046026; PID:g1046032; TIGR:MG329

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: GCC3

A:Start codon: GTG

C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

C:Keywords: duplication; GTP binding; nucleotide binding; P-loop

F:2-121/Domain: translation elongation factor Tu homology <ET1>

F:8-15/Region: nucleotide-binding motif A (P-loop)

F:118-121/Region: GTP-binding NKXD motif #status atypical

F:179-300/Domain: translation elongation factor Tu homology <ET2>

F:185-192/Region: nucleotide-binding motif A (P-loop)

F:297-300/Region: GTP-binding NKXD motif

F:332-334/Region: GTP-binding SAL/K motif #status atypical

Query Match 42.3%; Score 935.5; DB 2; Length 448;

Best Local Similarity 43.2%; Pred. No. 3.5e-50;

Matches 189; Conservative 86; Mismatches 154; Indels 9; Gaps 4;

QY 6 VAIVGRPVNGSTIFENRITVGERNSIVEDPQVTRDRIRYSGEWLTHFNIIIDGIEIGD 65

Db 4 VAIIGRNVGSTIFENRIQKPMIVSDPTPTTTRDFJGEGWKRRKRIAPIDRGLTAKQ 63

QY 66 APFOQIRAQAEIADADVIFMWANREGITQSDENVAQILYKSK-KPVLVANVKYDNM 124

Db 64 TPLOQLIALQQAALSOAKATIFELVLSIQEOLNSDPFYAKVLKKNKDKRPVILVYKNKENE 123

QY 125 EMRT---DYVDYSLGEGEPYPIISGSHGLGLGILLDAVYSH--FGEEEDDPYDEDTIRL 178

Db 124 NPKTAEEETIKDYSLGEGRPVYISAHGIGIGIDMLDLVKNQNLPRMEND--DLAKIRF 161

QY 179 SIIGRPVNGKSLVNALIGEDRVYSVNAAGTTIRAITDEYSYDGDYVLLIDTAGMRKKK 238

Db 182 CVIKRPVNGKSLINQVKNRVLVNSGEGTTIRAIIVPLVKNKEKFLIDTAGIKRKK 241

QY 239 VYESREKSVYRALKAIRSNVVLVIDAEGGIIIEODKRVAGVYHDEGKAVIIVYNNKMDT 298

Db 242 INMGIEIRASYIKTKLATIRSNVILIMWDGSKPISEODEVIGGLQAALLIPYIILVYNNKMDL 301

QY 299 VEKDSKTKKKEFEDEVRKEFOFLDYAQIAFYSAKERTRLTFLPYINEASENHKKRVQOST 358

Db 302 VLKNNNTNAVKKMLKHLHFKLHDEAPVLFISVLKNQRLNTEFEDKLIISOLETKVAATPL 361

QY 359 LNEVYTDIAISNPPPTDKGRLLNVFYATQVAIEPPTFVVVFVNDVLMHFSYKRYLENOIR 418

Db 362 LNDVYQQAQLNQPPLFGKRLQITVYANQTSQPHFVLFECNDKRYLHFSYARFLERKIR 421

```

OY      419  AAFGEECTPPIHIAARKN 436
Db      422  ENFGNSVPISLYFKSKN 439

RESULT 15
E91050
probable GTP-binding factor [Imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E91050
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurikawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A9629; MUID:21156231; PMID:11258796
A:Accession: E91050
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA836796.1; PID:g13362843; GSPDB:GND0154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
C:Gene: ECs3373
C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat

Query Match      41.7%; Score 923; DB 2; Length 503;
Best Local Similarity 42.2%; Pred. No. 2,4e-49;
Matches 198; Conservative 90; Mismatches 143; Indels 38; Gaps 4;

OY      4  PIVAVRPNVGKSTFNRIVGERSVIEDTPGTRDRIYSGEMLTHDPNIIDTGIEI 63
Db      16  PVALVGRPNVGSKTLNRLTRTRDALVAOPPGTLTRKRYGRAIEGREGTICIDTGIDG 75

OY      64  GDAPFOQIRAOEAIADADVILFMVNVREGIQTSDENVAOILYKSKRPVLAIVKVDN 123
Db      76  TEGDEVTRMAEQSLIAEADVLFVVDARAGLMPADEALAKHLRSEKPTFLVANKTDG 135

OY      124  MEMRTDYDPTSLGFGEPYPIYSGSHGAGLDLDLAVV----- 160
Db      136  LDPPQAAVDFVSLGLGIYPIAASHGKGVLSLEHVLPLPMEDLAPQEEVDEDAEYWAOF 195

OY      161  ---SHFGEEDDPYDEDT--IRLSIIRPNVGSKLVNALIGEDRVIVSWAGTTRDAID 215
Db      196  EAEENGGEEDDPDPQSLPIKLAIVGRPNVGSSTLNRILGGERVVYVIMPGTTRDSIT 255

OY      216  TEVSYDQDQVYLIDTAGMRKKGVYESTERYSVLRALKATERSNVVLVVIDAEGIIEDD 275
Db      256  IPMRDGQREVYLIDTAGVRRKGTDAVEKFSVYIKTLQALIEDANVVMVLVIDAREGSDQ 315

OY      276  KRVAGVYAHBOGKAVIYVKKMDYVERKSKTMKKFEDDEVRKEFOPLDYAOIAFVSAEKTR 335
Db      316  LSLGLFTLNSRKSIVIYVKNKMDGLSQKQVKEETDLCF--LDFIDFARVHFTSALHGSG 373

OY      336  LRTLEPTINAESENHKKRVQOSTLENVVTAISMNPTPKGRGLNVFVTAQVAIEPPTF 395
Db      374  VGNLFESYREAYIDSTSRVGTSMLTRITMAVEDHOPPLVGRGRVYIKATYAHAGGYNPPIY 433

OY      396  VVFNVDVLMHFSYKRYLQNIIRAFAFGECTPIHI-----IARKN 436
Db      434  VTHGNQVKDLPDYSKYRYLMTNFKSLDVMGSPRIQKKEGENPYANKRN 482

Search completed: September 10, 2003, 00:35:50
Job time : 44 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:19:46 ; Search time 23 Seconds
(without alignments)
891.463 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MTRPIVAIVGRPNVKGSTIF.....IRAFGEPTPIHIAKRRN 436

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	2213	100.0	436	1	ENG_A_STAM	099u15 staphylococ
2	1731	78.2	436	1	ENG_A_LISMO	08y5w8 listeria mo
3	1724	77.9	436	1	ENG_A_LISIN	092a71 listeria in
4	1699	76.8	436	1	ENG_A_BACU	P50743 bacillus su
5	1658	74.9	437	1	ENG_A_BACD	09kcd4 bacillus ha
6	1528	69.0	436	1	ENG_A_STRPY	09k1d0 streptococc
7	1523	68.8	436	1	ENG_A_STRP3	08k8j8 streptococc
8	1522	68.8	436	1	ENG_A_STRPN	097pc9 streptococc
9	1520	68.7	436	1	ENG_A_LACIA	09cjh6 lactococcus
10	1251.5	56.6	439	1	ENG_A_THETN	08t9j1 thermococ
11	1249.5	56.5	438	1	ENG_A_CLOPE	08t9j1 clostridium
12	1218.5	55.1	438	1	ENG_A_CLOAB	0971d7 clostridium
13	1178.5	53.3	440	1	ENG_A_FUSNN	08t9z7 fusobacteri
14	1037	46.9	453	1	ENG_A_ANASP	08t9z7 anabaena sp
15	1021	46.1	452	1	ENG_A_SYNY3	P74120 synechocyst
16	996	45.0	449	1	ENG_A_MYCPN	P73309 mycoplasma
17	995	45.0	442	1	ENG_A_UREPA	09t9q7 ureaplasma
18	984.5	44.5	435	1	ENG_A_MYCPU	09t9f1 mycoplasma
19	935.5	42.3	448	1	ENG_A_MYCCE	P47571 mycoplasma
20	923	41.7	490	1	ENG_A_ECOLI	P77254 escherichia
21	920	41.6	490	1	ENG_A_ECOL6	08t9f5 escherichia
22	910	41.1	490	1	ENG_A_SALTI	08t9p6 salmonella
23	910	41.1	490	1	ENG_A_SALTY	09xc18 salmonella
24	907.5	41.0	437	1	ENG_A_CHLTE	08t9k3 chlorobium
25	906.5	41.0	437	1	ENG_A_VIBCH	09t9w7 vibrio chol
26	898.5	40.6	496	1	ENG_A_VIBU	08t9f2 vibrio vuln
27	896.5	40.5	495	1	ENG_A_YERPE	08t9c2 yersinia pe
28	878.5	39.7	498	1	ENG_A_VIBPA	08t9s12 vibrio para
29	877.5	39.7	465	1	ENG_A_STRO	09t9w8 streptococ
30	864.5	39.1	433	1	ENG_A_AQUAE	06t9t49 aquifex aeo
31	862.5	39.0	519	1	ENG_A_CORCL	08t9k6 corynebacte
32	860.5	38.9	485	1	ENG_A_NEIMB	09t9z1 neisseria m
33	858.5	38.8	485	1	ENG_A_NEIMA	09t9v01 neisseria m

34	857.5	38.7	485	1	ENG_A_NEIGO	08t9a07 neisseria g
35	856	38.7	510	1	ENG_A_PASMU	P57812 pasteurella
36	852	38.5	504	1	ENG_A_HAEIN	P44536 haemophilus
37	846	38.2	447	1	ENG_A_RALSO	08y026 ralstonia s
38	842.5	38.1	439	1	ENG_A_THEMA	09t9f8 thermotoga
39	840.5	38.0	493	1	ENG_A_PSEAE	09t9j8 pseudomonas
40	808.5	36.5	479	1	ENG_A_RHITO	09t9d9 rhizobium 1
41	806.5	36.4	463	1	ENG_A_MYCU	03t912 mycobacteri
42	797.5	36.0	483	1	ENG_A_BRUME	08yfn2 brucella me
43	789	35.7	476	1	ENG_A_AGRF5	08t9d28 agrobacteri
44	784	35.4	438	1	ENG_A_DEIRA	09t9s19 delnoccocus
45	782.5	35.4	461	1	ENG_A_MYCLE	Q49884 mycobacteri

ALIGNMENTS

RESULT 1
ID ENG_A_STAM STANDARD: PRT: 436 AA.
AC 099u15:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENG_A OR SAV1475 OR SA1307 OR MM1364.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Haseyama A.,
RA Mutantani-Ul Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -!- SIMILARITY: BELONGS TO THE ERA/TREME FAMILY OF GTP-BINDING
CC PROTEINS. ENG_A SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AP003362; BAB57637.1; -
CC EMBL: AP003134; BAB42568.1; -
CC EMBL: AP004827; BAB95229.1; -
CC PIR: C89926; C89926.
CC HAMAP: MF_00195; -; 1.
CC InterPro: IPR005289; GTP-binding_dom.
CC InterPro: IPR006073; GTP1_OBG.

```

DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRPFAMS: TIGR00650; MG442; 2.
DR TIGRPFAMS: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 119 GTP 1 (POTENTIAL).
FT NP_BIND 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 GTP 2 (POTENTIAL).
FT NP_BIND 229 GTP 2 (POTENTIAL).
FT NP_BIND 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 48979 MW; 9EA4281B62F229D3 CRC64;

Query Match
Best local similarity 100.0%; Score 2213; DB 1; Length 436;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPPIVAIVGRPNVGKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTTHDFNIIDTGG 60
DB 1 MKPPIVAIVGRPNVGKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTTHDFNIIDTGG 60
QY 61 IETGAPFQTOIRAQAEIAIDEDAVIIFPVNVRREGTOSDEMAVAQILYKSKRPVLAIVK 120
DB 61 IETGAPFQTOIRAQAEIAIDEDAVIIFPVNVRREGTOSDEMAVAQILYKSKRPVLAIVK 120
QY 121 VNMEMRTVDYFYSLGFEPEYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
DB 121 VNMEMRTVDYFYSLGFEPEYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
QY 181 ITRPNVGKSSLVNAILGEDRVIVSNVAGTTTDAIDTYSYDQDYLITDAGMRKKGKYY 240
DB 181 ITRPNVGKSSLVNAILGEDRVIVSNVAGTTTDAIDTYSYDQDYLITDAGMRKKGKYY 240
QY 241 ESTEKSVLRALKAIERSNVVLVIDAEGGITEODKRVAGVHOGKAVIIVNKMDTVE 300
DB 241 ESTEKSVLRALKAIERSNVVLVIDAEGGITEODKRVAGVHOGKAVIIVNKMDTVE 300
QY 301 KDSKTMKKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPYINAESENHKKRVOSTLN 360
DB 301 KDSKTMKKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPYINAESENHKKRVOSTLN 360
QY 361 EYVTDASNNPTPTDKGRNLNFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
DB 361 EYVTDASNNPTPTDKGRNLNFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
QY 421 FGEETGPIHIIARKRN 436
DB 421 FGEETGPIHIIARKRN 436

RESULT 2
ENGA_LISMO STANDARD; PRT; 436 AA.
AC 0815M8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA OR LMO1937.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-se / Setovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Franjeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baguerio F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Entlan K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

```

```

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Klett J., Kuhn M., Kunst F., Kurapkai G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomes of Listeria species."
RL Science 294:849-852(2001).
CC -i- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AL591981; CAD00015.1; -.
DR PIR: A11316; A11316.
DR Listlist: LMO01937; -.
DR HAMAP: ME_00195; -.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRPFAMS: TIGR00650; MG442; 2.
DR TIGRPFAMS: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 119 GTP 1 (POTENTIAL).
FT NP_BIND 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 GTP 2 (POTENTIAL).
FT NP_BIND 229 GTP 2 (POTENTIAL).
FT NP_BIND 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 49144 MW; 5B435C8701AE6A3 CRC64;

Query Match
Best local similarity 78.2%; Score 1731; DB 1; Length 436;
Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

QY 1 MKPPIVAIVGRPNVGKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTTHDFNIIDTGG 60
DB 1 MKPPIVAIVGRPNVGKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTTHDFNIIDTGG 60
QY 61 IETGAPFQTOIRAQAEIAIDEDAVIIFPVNVRREGTOSDEMAVAQILYKSKRPVLAIVK 120
DB 61 IDSDEPLEQJIRAQAEIAIDEDAVIIFPVNVRREGTOSDEMAVAQILYKSKRPVLAIVK 120
QY 121 VNMEMRTVDYFYSLGFEPEYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
DB 121 VNMEMRTVDYFYSLGFEPEYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
QY 181 ITRPNVGKSSLVNAILGEDRVIVSNVAGTTTDAIDTYSYDQDYLITDAGMRKKGKYY 240
DB 181 ITRPNVGKSSLVNAILGEDRVIVSNVAGTTTDAIDTYSYDQDYLITDAGMRKKGKYY 240
QY 241 ESTEKSVLRALKAIERSNVVLVIDAEGGITEODKRVAGVHOGKAVIIVNKMDTVE 300
DB 241 ESTEKSVLRALKAIERSNVVLVIDAEGGITEODKRVAGVHOGKAVIIVNKMDTVE 300
QY 301 KDSKTMKKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPYINAESENHKKRVOSTLN 360
DB 301 KDSKTMKKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPYINAESENHKKRVOSTLN 360
QY 361 EYVTDASNNPTPTDKGRNLNFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
DB 361 DIYSADVANNPSPMDKGRILKIFTYQVAIVKPTFVFNVDLMHFSYERLENIIRAA 420
QY 421 FGEETGPIHIIARKRN 435

```

```

Db          421  PFEGCTPIRVIAKR 435
          | | | | | | | | | |
RESULT 3
ENGA_LISIN
ID  ENGA_LISIN  STANDARD:  PRT:  436 AA.
AC  Q92A71;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Probable GTP-binding protein enga.
GN  ENGA OR LIN2051.
OS  Listeria innocua.
OC  Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX  NCBI_TaxID=1642;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CLIP 11262 / Serovar 6a;
RX  MEDLINE=21537279; PubMed=11679669;
RA  Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA  Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA  Chabdit A., Chetouni F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA  Entian K.-D., Esli H., Garcia-del Portillo F., Garrido P.,
RA  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA  Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkhat G.,
RA  Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA  Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Puceat R.,
RA  Remmel B., Rose M., Schueter T., Simoes N., Tietze A.,
RA  Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RT  "Comparative genomics of Listeria species.";
RL  Science 294:849-852(2001).
CC  -1 SIMILARITY: BELONGS TO THE ERA/TRNE FAMILY OF GTP-BINDING
CC  PROTEINS. ENGA SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: AL596170; CAC97281.1; -
DR  PIR: A11688; A11688
DR  ListList: LIN2051; -
DR  HAMAP: MF_00195; -; 1.
DR  InterPro: IPR005289; GTP-binding_dom.
DR  InterPro: IPR006073; GTP_OBG.
DR  InterPro: IPR002917; MMR_HSR1.
DR  InterPro: IPR005225; Small_GTP.
DR  Pfam: PF01926; MMR_HSR1. 1.
DR  PRINTS: PR00326; GTP_OBG.
DR  TIGRFAMs: TIGR00650; MG442; 2.
DR  TIGRFAMs: TIGR00231; small_GTP. 2.
KW  GTP-binding; Repeat; Complete proteome.
FT  NP_BIND 10 17 GTP 1 (POTENTIAL).
FT  NP_BIND 57 61 GTP 1 (POTENTIAL).
FT  NP_BIND 119 122 GTP 2 (POTENTIAL).
FT  NP_BIND 182 189 GTP 2 (POTENTIAL).
FT  NP_BIND 229 233 GTP 2 (POTENTIAL).
FT  NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ  SEQUENCE 436 AA; 49144 MW; 1A435C1970EB66C8 CRC64;
Query Match 77.9%; Score 1724; DB 1; Length 436;
Best Local Similarity 73.1%; Pred. No. 2e-98;
Matches 318; Conservative 67; Mismatches 50; Indels 0; Gaps 0;

```

```

Oy          61  IEIGDAPFQOTIRAQAEIAIDEDAVIIFPMVNVREGTOSDENVAQILKSKKPVAVNK 120
          | | | | | | | | | |
Db          61  IDLSDEPFLEQIRAAQAEIAIDEDAVIIFITNREGVTDADQVAKILVRSNKPVLAINK 120
Oy          121  VDNMEKRDVDYDFYSLGCEPPIPGISGIGLGLDLDLVNVSHFGGEEDPYDEDIIRLSI 180
          | | | | | | | | | |
Db          121  VDNPEMRQOIIDFYSLGCEPPIPGISGIGLGLDLDLVNRAHFPKEEEDPYDEDIIRLSI 180
Oy          181  IGRPNVGSLSVALNALLGEDRVVSNVAGTTRPAIDTEXSYDQODVLDITAGMRKRGKY 240
          | | | | | | | | | |
Db          181  IGRPNVGSLSVALNALLGEDRVVSNVAGTTRPAIDTEXSYDQODVLDITAGMRKRGKY 240
Oy          241  ESTERYSVLRALKAIERSNVVLVIDAEGCIEQDKRVAAGAHQEGKAVIYVNVKWDVE 300
          | | | | | | | | | |
Db          241  ESTERYSVLRALKAIERSNVVLVIDAEGCIEQDKRVAAGAHQEGKAVIYVNVKWDVE 300
Oy          301  KDSKTMKRFEDVREKPEQFLDYAQAIPASAKERTLRTLPFINAASEMHNKRVQSTLN 360
          | | | | | | | | | |
Db          301  KDKETINWTEIDIRQFQFLSYAPIVFVSAKTRKRLNLEPLINQVSDNHSLSRVQSMUN 360
Oy          361  EYVTDIAISNPTPTDKGRNLNVEYATQVAVIEPPTFVFNVDVLMHFYSKRYLEQIRAA 420
          | | | | | | | | | |
Db          361  DIVISDAVAMNPSPMQKGRKLFYTTQVAVKRPTEVFNVDVLMHFYSKRYLEQIRAA 420
Oy          421  FGSECTPIHIAKR 435
          | | | | | | | | | |
Db          421  PFEGCTPIRVIAKR 435
          | | | | | | | | | |
RESULT 4
ENGA_BACSU
ID  ENGA_BACSU  STANDARD:  PRT:  436 AA.
AC  P50743;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Probable GTP-binding protein enga.
GN  ENGA.
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX  NCBI_TaxID=1423;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=168 / Marburg;
RX  MEDLINE=96349105; PubMed=8760912;
RA  Sorokin A.V., Azavedo V., Zumbstein E., Galleron N., Ehrlrich S.D.,
RA  Serror P.;
RT  "Sequence analysis of the Bacillus subtilis chromosome region between
RT  the sera and kds loci cloned in a yeast artificial chromosome.";
RL  Microbiology 142:2005-2016(1996).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=168;
RX  MEDLINE=98044033; PubMed=9384377;
RA  Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA  Azevedo V., Bertero M.G., Bessieres P., Boloquin A., Borchert S.,
RA  Borries R., Bourasier L., Brans A., Braun M., Brignelli S.C., Bron S.,
RA  Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,
RA  Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA  Denizot F., Devine K.M., Dusterhoft A., Ehrlrich S.D., Emerson P.T.,
RA  Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA  Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA  Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grand G.,
RA  Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA  Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA  Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA  Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
RA  Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA  Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA  Medina N., Mellado R.P., Mizuno M., Mestl D., Nakai S., Noback M.,
RA  Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA  Paro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
RA  Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

```

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield E.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A.,
 RA Viari A., Wambolt R., Wedler E., Wedler H., Wetschegg T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE OF 185-436 FROM N.A.
 RC STRAIN-168;
 RX MEDLINE=96011379; PubMed=7592341;
 RA Moridoni H.R., de Mendoza D., Cronan J.E. Jr.;
 RT "Synthesis of sn-glycerol 3-phosphate, a key precursor of membrane
 RT lipids, in *Bacillus subtilis*.";
 RL J. Bacteriol. 177:5899-5903(1995).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENGA SUBFAMILY.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
 CC frameshift.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L47648; AAC83966.1; -
 DR EMBL: U32164; CAB14200.1; -
 DR PIR: A69936; A69936.
 DR Subtilist; BG11443; enga.
 DR HAMAP: MF_00195; -; 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small-GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR PRINTS: PR00326; GTP_OBG.
 DR TIGRFAWS: TIGR00650; MG442; 2.
 DR TIGRFAWS: TIGR00231; small_GTP; 2.
 KM GTP-binding: Repeat; Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 182 189 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 FT CONFLICT 186 187 VG -> CR (IN REF. 2).
 SQ SEQUENCE 436 AA; 48769 MM; A5CC7028F8B5A442 CRC64;

Query Match 76.8%; Score 1699; DB 1; Length 436;
 Best Local Similarity 73.6%; Pred. No. 6.8e-97;
 Matches 320; Conservative 57; Mismatches 58; Indels 0; Gaps 0;

QY 181 IGRPNWKKSLVNAIIGEDRYIVSNAGTTROAIDREYSYDGDVYLIDTAGMKKKGY 240
 DB 181 IGRPNWKKSLVNAIIGEDRYIVSNAGTTROAIDREYSYDGDVYLIDTAGMKKKGY 240
 QY 241 ESTREKSYVALKAIRSNVNLVVADEGIIKODRVRVGYAEHOGKAVIYVKKNDYVE 300
 DB 241 ETTEKISYVALKAIRSNVNLVVADEGIIKODRVRVGYAEHOGKAVIYVKKNDYVE 300
 QY 301 KDSKTKKEDEVRKEFOFLDYAQAIVSAKERTRLTLEPIYNEASENKKRVSSTIN 360
 DB 301 KDSKTKKEDEVRKEFOFLDYAQAIVSAKERTRLTLEPIYNEASENKKRVSSTIN 360
 QY 361 EYVTDALISNPTPTDGRNLNVYATQVAIEPTFPVYVNDVLMFSTKRYLENDIRA 420
 DB 361 DVTMDAVANMPRTTHNGSRUKIYATOVSKPPSFVAVNDPELMHFSEYERLENIRDA 420
 QY 421 FGEGTPIHIAKR 435
 DB 421 FGEGTPIHIAKR 435

RESULT 5
 ENGA_BACHD STANDARD; PRT; 437 AA.
 ID ENGA_BACHD
 AC Q9KCD4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable GTP-binding protein enga.
 GN ENGA OR BH1638
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 ON NCBI_TaxID=86665;
 RX NCBI_TaxID=86665;
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512562; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENGA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP001512; BAB05357.1; -
 DR PIR: F83854; F83854.
 DR HAMAP: MF_00195; -; 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small-GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR PRINTS: PR00326; GTP_OBG.
 DR TIGRFAWS: TIGR00650; MG442; 2.
 DR TIGRFAWS: TIGR00231; small_GTP; 2.
 KM GTP-binding: Repeat; Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 182 189 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 SQ SEQUENCE 437 AA; 49024 MM; C1A48D55A39AD4B CRC64;

	Query Match	74.9%.	Score 1658:	DB 1:	Length 437:
	Best Local Similarity	70.9%.	Pred. No. 2.2e-94:		
	Matches 309;	Conservative 67;	Mismatches 60;	Indels 0;	Gaps 0;
QY	1 MTKPDIIVAVIGRPNGKSTIFNRIRIGERSIYIEDPGVTRDRIRYSGEWMLTHFDNIIDGG 60				
Dd	1 MSKVIAIVIGRPNPKSTIFFRKIRYGERAIYVEDRGVTDRIRISGEMLNREFNVITDGG 60				
QY	121 VDNMEMRPDYDFEFLSGCEPYRPSGSHGLDGLDLDAVNSHFGSEEDDPVDEDIRLSL 180				
Dd	121 IDHPDMDELEEFYSLGIGDPIPLISGAGLGSLGDLDCVEHFPEPDEDDDVDDEDIRLSL 180				
QY	181 IGRPNVGSSLVNAILCEDRYIVSNVATGRDAITDEXSYDGQDVLLIDTGA MRKKGY 240				
Dd	181 IGRPNVGSSLVNAMLGERRIVSNSIPTRTDAIDTAISRDDQEVLLIDTAGMRKKGY 240				
QY	241 ESTEKYSYLRLAKAIERSNVVLVVIDAEGIIEDDKRVAGSAHEOGKAVIIVVNKMDTVE 300				
Dd	241 KDSFTMMKFEEDEVKKEPOFLDYAOAFASAKERTLRLLFRPINASENNHKRRVOSTLN 360				
QY	361 EYVTDALISMNPTRPDKGRLVWFAYATQVAIEPPFPFVVNVNDLMHFSKYRLNQIRAA 420				
Dd	361 DLVMDAAVMNPTPIDHGKRKLINVTQVAVGPRTTFVFVVNDPELMHFSARPLENRADT 420				
QY	421 FFEGETPIHIITARRKN 436				
Dd	421 FEFEGETPIKIITARRKN 436				
RESULT 6					
ENGA_STRPY	STANDARD:	PRT:	436 AA.		
AC	Q9A1D0:				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Probable GTP-binding protein enga.				
GN	ENGA OR PGDA OR SPY0341 OR SPM18_0334.				
OS	Streptococcus pyogenes, and				
OS	Streptococcus pyogenes (serotype M18).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
CC	Streptococcus				
OX	NCBI_TaxID=1314, 186103;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SF370 / ATCC 700294 / Serotype M1;				
RX	MEDLINE=2192684; PubMed=112962296;				
RA	Ferretti J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,				
RA	Primeaux C., Sezate S., Stuvorov A.N., Kenton S., Lai H.S., Lin S.P.,				
RA	Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,				
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;				
RT	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).				
RJ	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-MGAS8232 / Serotype M18;				
RX	MEDLINE=21927593; PubMed=11917108;				
RA	Smoot J.C., Barbant K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,				
RA	Slyva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,				
RA	Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,				
RA	Kapur V., Daly J.A., Veasy L.G., Musser J.K.;				
RT	"Genome sequence and comparability microarray analysis of serotype M18				
RT	group A Streptococcus strains associated with acute rheumatic fever				
RT	outbreaks.";				

RL	Proc. Acad. Sci. U.S.A. 99:4668-4673(2002).	
CC	-1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING	
CC	PROTEINS. ENGA SUBFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	CC the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AE006498; AAK33393.1; -	
DR	EMBL; AE009978; AAL97088.1; -	
DR	HAMAP; MF_00195; -; 1	
DR	InterPro; IPR005289; GTP-binding_dom.	
DR	InterPro; IPR006073; GTP1_OBG.	
DR	InterPro; IPR002917; MMR_HSR1.	
DR	InterPro; IPR005325; Small-GTP.	
DR	Pfam; PF01926; MMR_HSR1; 1.	
DR	PRINTS; PR00326; GTP1OBG.	
DR	TIGRFAMS; TIGR00650; MG442.2	
DR	TIGRFAMS; TIGR00231; small_GTP; 2.	
KW	GTP-binding; Repeat: Complete proteome.	
FT	NP_BIND 10	GTP 1 (POTENTIAL).
FT	NP_BIND 57	GTP 1 (POTENTIAL).
FT	NP_BIND 119	GTP 1 (POTENTIAL).
FT	NP_BIND 181	GTP 2 (POTENTIAL).
FT	NP_BIND 229	GTP 2 (POTENTIAL).
FT	NP_BIND 294	GTP 2 (POTENTIAL).
QO	SEQUENCE 436 AA; 48801 MW; EDEDGDI51675323A CRC64;	

Query Match	69.0%	Score 1528	DB 1	Length 436
Best Local Similarity	66.1%	Pred. No. 1.9e-86		
Matches 288	Conservative 75	Mismatches 71	Indels 2	Gaps 2
QY	1	MTKEPIVAVIGRPNGKSTIFNRIGVERSVIEDTPGVYTRDRISYSGEWLTHDFNIIDTGG	60	
Db	1	MVLPTVAIVIGRPNGKSTLFLRINGERISYIEDVEGYTRDRIRYATGEMLNQFSIIDTG	60	
QY	61	TEIGDAPPQOIRKQAEIADDEADVIIFMNVRREGLTQSDENMAOILYKSKPPVLAVNK	120	
Db	61	IDVDAPFMEQIKHQAOIAMEADVIYFVVGSKGVDADAEYSKILRTYTPVILA VNK	120	
QY	121	VDNMEMRDVYDFSLGFGPEYPISSGSLGGLGLLAAVSHFCEBEDPYDEDTIRLSI	180	
Db	121	VDNEMERNDITDFSLGLGDPYPVSSVHGISTGVDLAIYVNLVEEAEEND-DIIRSL	179	
QY	181	IGRPNGKSSLVNAILGEDRVIVSNVAGTTSDAIDTETYS-YDGDDYVLIDTAGNRKKGV	239	
Db	180	IGRPNGKSSLINMILGEDRVIASPVAGTTSDAIDTHTDADGGEFTMIDTAGNRKSKGI	239	
QY	240	YESIEKTSVLAALAIERSNNVLYVIDAEQIIEODKRVAGIAHEOGCAVYIVVANKDVT	299	
Db	240	YENIEKTSVYMAMARAIDSDVLLVINAEDEIREDKRIAGFAHEAGCMIIIVANKDPTI	299	
QY	300	EKDSKTKMKFEDERYKEFOFLDYQOIAFVASKERTRLTLETPPYINESENHKRRVOSTL	359	
Db	300	DKDHTYAKWEADRIDQOFLTYAIPITFVSALTYQRLNKLEDDLKRISESQNKIIPSAVL	359	
QY	360	NEVYTDAISNPTPTDKGRNLNVETAOVAIEPTFVVVNDVDELHHSYKRYLENDQIRA	419	
Db	360	NDVIMDAIINPTPTDKGRKRIKIFYATQSVSKPPTFVVVFNVEEELMHFSYLRPLENQIRA	419	
QY	420	AFGEGETPIHILIAKKR 435		
Db	420	AFTEGETPIHILIAKKR 435		

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR PGDA OR SPY3_0249 OR SP51610.
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=198466;
 RX NCBI_TaxID=198466;
 RP SEQUENCE FROM N.A.
 RC STRAIN-MGAS315 / Serotype M3;
 RX MEDLINE=2213808; PubMed=1212206;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "genome sequence of a serotype M3 strain of group A Streptococcus;
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
 RA Hayaishi H., Hamada S.;
 RT "The genome of invasive Streptococcus pyogenes: a comparative analysis
 RT of S. pyogenes SSI-1, SPI70 and MGAS8232.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE014141; AAM78856.1; -
 DR EMBL: AP005146; BAC64705.1; -
 DR HAMAP: MF_00195; -; 1.
 DR Interpro: IPR005289; GTP-binding_dom.
 DR Interpro: IPR006073; GTP1_OBG.
 DR Interpro: IPR002917; MMR_HSR1.
 DR Interpro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR PRINTS: PR00326; GTP1_OBG.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR TIGRFAMs: TIGR00231; small_GTP; 2.
 KM GTP-binding: Repeat; Complete proteome.
 FT NP_BIND 10 17 GTP 1 (POTENTIAL).
 FT NP_BIND 57 61 GTP 1 (POTENTIAL).
 FT NP_BIND 119 122 GTP 1 (POTENTIAL).
 FT NP_BIND 181 188 GTP 2 (POTENTIAL).
 FT NP_BIND 229 233 GTP 2 (POTENTIAL).
 FT NP_BIND 294 297 GTP 2 (POTENTIAL).
 SQ SEQUENCE 436 AA; 48771 MW; F2DFCD5B674323A CRC64;
 Query Match 68.88; Score 1523; DB 1; Length 436;
 Best Local Similarly 65.88; Pred. No. 3.9e-86;
 Matches 287; Conservative 75; Mismatches 72; Indels 2; Gaps 2;
 QY 1 MFKPIAIVGRPNVNGSTINRIYGEVSTVEDTPGVTRDRITVSSSGEMLTHDFIIDTGG 60
 DB 1 MVLPTVAIVGRPNVNGSTINRIYGEVSTVEDTPGVTRDRITVSSSGEMLTHDFIIDTGG 60
 QY 61 IEIGDAPFQIIRAQAEIAIDADVILTFNVREGLTQSGEMVAQIIKSKKPPVLAIVNK 120
 DB 61 IDVDNAPFMQIKHQADIAEADVIVFYVSGKEGYDADDEYVSKILRTINTPTIIVLVNK 120

QY 121 VDNEMKRTDYDYSIGFCEPPIYSGSHGICGLDLDVAVSHFCEEDPDYEDTIRLSI 180
 DB 121 VDNEMKRTDYDYSIGFCEPPIYSGSHGICGLDLDVAVSHFCEEDPDYEDTIRLSI 179
 QY 181 IGRPNVNGKSLVNAIIGEDVIVSNVAGTTRDAIDPEYS-YDGDVYLDITAGMRKGV 239
 DB 180 IGRPNVNGKSLVNAIIGEDVIVSNVAGTTRDAIDPEYS-YDGDVYLDITAGMRKGV 239
 QY 240 YESTEKYSVALKATERSNVVLVIDAEGIIEDOKRVAGVAHEQGAIVVNRMDTV 299
 DB 240 YENTEKYSVNRARAIADRSQVLMVINEEGINEYKRIAGFHAEGKMIIVNKMDAI 299
 QY 300 EKDSKTKKPEDEYKKEFQFLDYAQAIFASAKERTLRITFPYINAESEHKKRVOSSTL 359
 DB 300 DKDNHVTAKMEADIRDOFQFLTYAPITFVSALTKORLKLPLDIKRISSESONKRIPSAYL 359
 QY 360 NEVYTAISMNPTPTGKRLNPFYATQVAIEPTFPVFNVDVLMHFSKRYLENDIRA 419
 DB 360 NDVIMDAIINPTPTDGKRLKIFATQVSVKPPTEVFNBEELMHFSTRLFNQIRA 419
 QY 420 AFGEFGPIHIIARRK 435
 DB 420 AFTEFGPIHIIARRK 435
 RESULT 8
 ENG A STRPN STANDARD; PRT; 436 AA.
 ID ENG A STRPN
 AC 097PC9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR SERA OR SPI709 OR SPI1553.
 OS Streptococcus pneumoniae, and
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313, 171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapfel E., Khouri H., Wolf A.M., Ufferbach T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC BAA-255 / R6;
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller R.E., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McHenry S.M., McHenry M., Mclester K., Mundy C.W., Nicot T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.T.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR1 EMBL: AE007464; AAK75787.1; -
DR EMBL: AE008523; AAL00357.1; -
DR PIR: B95199; B95199.
DR PIR: H98065; H98065.
DR TIGR: SP1709; -
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRFAMS: TIGR00650; MG442; 2.
DR TIGRFAMS: TIGR00231; small_gtp; 2.
KM GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 181 188 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 49082 MW; B013890D5285BBD CRC64;

Query Match 68.8%; Score 1522; DB 1; Length 436;
Best Local Similarity 66.2%; Pred. No. 4.4e-86;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIYGERYSIYEDPGVTRDRYSSGEMLTHTDFTIDTGG 60
DB 1 MALPTVAIVGRPNVSKSTIFNRIYGERYSIYEDPGVTRDRYSSGEMLTHTDFTIDTGG 60
QY 61 IEIGDAPFOTQIRAOAEIAIDEDVYIIFMVNVRGLTQSDPMVAOILYKSKRPVLAIVNK 120
DB 61 IDVDAPFMEDIKKQAEIAMEADYIVVSGKEGIDTDADEVYARKLTKTKRPVLAIVNK 120
QY 121 VDNEMKRTDYDFYSLGFGEPYPISSGSHGLGLDVAIVSHFGE--EEDDPYEDDTIRL 178
DB 121 VDNEMKRTDYDFYSLGFGEPYPISSGSHGLGLDVAIVSHFGE--EEDDPYEDDTIRL 178
QY 179 SIIGRPVNGKSLVNAIIGEDRVYVSNAGTTRADITREYS-YDGDYVLLIDTAGMRKG 237
DB 178 SLIGRPVNGKSLVNAIIGEDRVYVSNAGTTRADITREYS-YDGDYVLLIDTAGMRKG 237
QY 238 KYVESTKYSVLRAIKAIERSNVLVVIDAEOGIIEODKRVAGYAHEGKAVVIVVNMKD 297
DB 238 KYVENTEKYSVMRAIKAIERSNVLVVIDAEOGIIEODKRVAGYAHEGKAVVIVVNMKD 297
QY 298 TLEKDNHMKWEEDIREQFQYLPYAPIIFVSALTQKHLKHPKIKISSQNRIISA 357
DB 298 TLEKDNHMKWEEDIREQFQYLPYAPIIFVSALTQKHLKHPKIKISSQNRIISA 357
QY 358 TLNEVYDAISMNFTPDGKRLNVFYATQVAIEPPTFVVVNDVLMHESYKRYLENOI 417
DB 358 VLNVDVDAISMNFTPDGKRLNVFYATQVAIEPPTFVVVNDVLMHESYKRYLENOI 417
QY 418 RAAGFEGEPTIHTAKRR 435
DB 418 RKAFFEGEPTIHTAKRR 435

RESULT 9
ID ENCA_LACIA STANDARD: PRT: 436 AA.
AC Q9CHH6: 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enga.
GN ENCA OR YPHL OR IL0755.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403.
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winkler P., Manger S., Jallion O., Malarne K.,
RA Welzenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RL Lactis ssp. Lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -I- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENCA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR1 EMBL: AE006309; AAK04853.1; -
DR PIR: C86719; C86719.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRFAMS: TIGR00650; MG442; 2.
DR TIGRFAMS: TIGR00231; small_gtp; 2.
KM GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 181 188 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 48866 MW; 1D9D0081F9865B9D CRC64;

Query Match 68.7%; Score 1520; DB 1; Length 436;
Best Local Similarity 66.3%; Pred. No. 5.9e-86;
Matches 289; Conservative 69; Mismatches 76; Indels 2; Gaps 2;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIYGERYSIYEDPGVTRDRYSSGEMLTHTDFTIDTGG 60
DB 1 MSLPTVAIVGRPNVSKSTIFNRIYGERYSIYEDPGVTRDRYSSGEMLTHTDFTIDTGG 60
QY 61 IEIGDAPFOTQIRAOAEIAIDEDVYIIFMVNVRGLTQSDPMVAOILYKSKRPVLAIVNK 120
DB 61 IELSDPEPMTIELRAOAEIAMEADYIVAVSGEIGITDADEVYARKLTKTKRPVLAIVNK 120
QY 121 VDNEMKRTDYDFYSLGFGEPYPISSGSHGLGLDVAIVSHFGEEDDPYEDDTIRLSI 180
DB 121 VDNEMKRTDYDFYSLGFGEPYPISSGSHGLGLDVAIVSHFGEEDDPYEDDTIRLSI 180
QY 181 IGRPNVNGKSLVNAIIGEDRVYVSNAGTTRADITREYS-YDGDYVLLIDTAGMRKGK 239
DB 180 IGRPNVNGKSLVNAIIGEDRVYVSNAGTTRADITREYS-YDGDYVLLIDTAGMRKGK 239
QY 240 YESTEKYSVLRAIKAIERSNVLVVIDAEOGIIEODKRVAGYAHEGKAVVIVVNMKD 299
DB 240 YENTEKYSVMRAIKAIERSNVLVVIDAEOGIIEODKRVAGYAHEGKAVVIVVNMKD 299
QY 300 EKDSKTKMKFEDEVKKEFQPLDYAOIAIVSAKERTRLTLPYINASENHKKRVOSTL 359
DB 300 EKDNMTKMKFELEITRKKEFLDYAPIVVSAAKGTGRMLKPLDMIKIHHOANLRISSSVL 359

QY	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 120 123 GTP 1 (POTENTIAL).
FT NP_BIND 183 190 GTP 2 (POTENTIAL).
FT NP_BIND 230 234 GTP 2 (POTENTIAL).
FT NP_BIND 295 298 GTP 2 (POTENTIAL).
SQ SEQUENCE 438 AA; 49586 MW; 9658B0BC98398396 CRC64;

Query Match
Best Local Similarity 53.4%; Pred. No. 2e-69;
Matches 233; Conservative 92; Mismatches 110; Indels 1; Gaps 1;

QY 1 MTKPIVAIGRPNVGKSTIFNRIIGERSIVDETPGVTRDRISYSGEWLTHDPNIIDTG 60
1 MSKPIVAVNGRPNVGKSTIFNKLAKKRISIVDTGCVTRDRYAESEWLNKRFMTIDTG 60
DB 1 IEGDAP-FQTOIRAOAETIADEADVIIFMVNVRGLTQSDENVAQILKSKRPVLA 119
61 IEGDAP-FQTOIRAOAETIADEADVIIFMVNVRGLTQSDENVAQILKSKRPVLA 119
DB 1 IEPESSDILVKOMRROAOIAIEMADVIIFVVDGKEGLTAADQEVAKMLKSKRPVLA 120
61 IEPESSDILVKOMRROAOIAIEMADVIIFVVDGKEGLTAADQEVAKMLKSKRPVLA 120
QY 120 KYDNMEMRTDVFYSLGGEPRYPSGSHGLGLDLDAVSHFGEEDPYDEDTIRLS 179
121 KIDRLALENSYEFLNGLGDPITISASGLGLGMLDEVKXFPNDPSEDEDEYIRTA 180
DB 121 KIDRLALENSYEFLNGLGDPITISASGLGLGMLDEVKXFPNDPSEDEDEYIRTA 180
QY 180 IIGRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 239
181 MGKRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 240
DB 181 MGKRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 240
QY 240 YESTEKYSVLRAKAIERSNVVLYVDAEGIIEDOKRVAAGAHGKRAVIVVNMKV 299
241 KEIERYSVIRTYAIEKADVALIVDAEOGITDEODEKITGAHEKNKAILVYVNMKMDLI 300
DB 241 KEIERYSVIRTYAIEKADVALIVDAEOGITDEODEKITGAHEKNKAILVYVNMKMDLI 300
QY 300 EKDSKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 359
301 EKDDKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 360
DB 301 EKDDKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 360
QY 360 NEVYDAISMNPTPDKGRRLNVFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 419
361 NDVISAVALMKRPPVVALKRLKIYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 420
DB 361 NDVISAVALMKRPPVVALKRLKIYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 420
QY 420 AFGECTPIHITARRK 435
421 SFDEGTGIMTKFER 436
DB 421 SFDEGTGIMTKFER 436

RESULT 12
ENGA_CLOAB STANDARD; PRT; 438 AA.
ID ENGA_CLOAB
AC 097ID7;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA OR CAC1711.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
CC Clostridium.
OX NCBI_TaxID=1488;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
Gibson R.L., Lee H.M., Dupois J., Qiu D., Hittl J., Wolf Y.T.,
Tatsov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE007680; AKK79677.1; -
DR PIR: B97111; B97111.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR PRINTS: PR00449; RASTRNSFRMG.
DR TIGRFS: TIGR00650; MG442; 2.
DR TIGRFS: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 120 123 GTP 1 (POTENTIAL).
FT NP_BIND 183 190 GTP 2 (POTENTIAL).
FT NP_BIND 230 234 GTP 2 (POTENTIAL).
FT NP_BIND 295 298 GTP 2 (POTENTIAL).
SQ SEQUENCE 438 AA; 49797 MW; 7765735BD7355C7 CRC64;

Query Match
Best Local Similarity 52.8%; Score 1218.5; DB 1; Length 438;
Matches 230; Conservative 88; Mismatches 117; Indels 1; Gaps 1;

QY 1 MTKPIVAIGRPNVGKSTIFNRIIGERSIVDETPGVTRDRISYSGEWLTHDPNIIDTG 60
1 MSKPIVAVNGRPNVGKSTIFNKLAKKRISIVDTGCVTRDRYAESEWLNKRFMTIDTG 60
DB 1 IEGDAP-FQTOIRAOAETIADEADVIIFMVNVRGLTQSDENVAQILKSKRPVLA 119
61 IEGDAP-FQTOIRAOAETIADEADVIIFMVNVRGLTQSDENVAQILKSKRPVLA 119
DB 61 IEGDAP-FQTOIRAOAETIADEADVIIFMVNVRGLTQSDENVAQILKSKRPVLA 119
QY 120 KYDNMEMRTDVFYSLGGEPRYPSGSHGLGLDLDAVSHFGEEDPYDEDTIRLS 179
121 KIDRLALENSYEFLNGLGDPITISASGLGLGMLDEVKXFPNDPSEDEDEYIRTA 180
DB 121 KIDRLALENSYEFLNGLGDPITISASGLGLGMLDEVKXFPNDPSEDEDEYIRTA 180
QY 180 IIGRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 239
181 MGKRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 240
DB 181 MGKRPNVGKSLVNAALGEDRVIVSNVAGTTRDAIDTEXSYDGODVILDTAGMRKGV 240
QY 240 YESTEKYSVLRAKAIERSNVVLYVDAEGIIEDOKRVAAGAHGKRAVIVVNMKV 299
241 KEIERYSVIRTYAIEKADVALIVDAEOGITDEODEKITGAHEKNKAILVYVNMKMDLI 300
DB 241 KEIERYSVIRTYAIEKADVALIVDAEOGITDEODEKITGAHEKNKAILVYVNMKMDLI 300
QY 300 EKDSKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 359
301 EKDDKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 360
DB 301 EKDDKTMKFEDEVERKEFOFLDYAOIAPYSAKERTRLRFLFYINASENHKKRVOSTL 360
QY 360 NEVYDAISMNPTPDKGRRLNVFYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 419
361 NDVISAVALMKRPPVVALKRLKIYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 420
DB 361 NDVISAVALMKRPPVVALKRLKIYATQVAIEPTFFVFNVDVLMHFSYKRYLENOIRA 420
QY 420 AFGECTPIHITARRK 435
421 SFDEGTGIMTKFER 436
DB 421 SFDEGTGIMTKFER 436

RESULT 13
ENGA_FUSNN STANDARD; PRT; 440 AA.
ID ENGA_FUSNN
AC 08RGV7;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR FN0170.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 CC Fusobacterium.
 OX NCBI_TaxId=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RA MEDLINE=21886394; PubMed=11889109;
 RA Kapralov V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhatnagar A., Barman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Feinstein M., Kyriides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium *Fusobacterium*
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE010530; AAL94376.1; -
 DR HAMAP: MF_00195; -; 1
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR TIGRFAMs: TIGR00231; small_GTP; 2.
 KM GTP-binding; Repeat; Complete proteome.
 FT NP-BIND 9 16 GTP 1 (POTENTIAL).
 FT NP-BIND 56 60 GTP 1 (POTENTIAL).
 FT NP-BIND 119 122 GTP 1 (POTENTIAL).
 FT NP-BIND 183 190 GTP 2 (POTENTIAL).
 FT NP-BIND 230 234 GTP 2 (POTENTIAL).
 FT NP-BIND 295 298 GTP 2 (POTENTIAL).
 SQ SEQUENCE 440 AA: 49608 MW: BEC805F839BE2A05 CRC64;
 Query Match 53.3%; Score 1178.5; DB 1; Length 440;
 Best Local Similarity 53.6%; Pred. No. 43e-65;
 Matches 221; Conservative 81; Mismatches 116; Indels 3; Gaps 3;
 QY 3 KPIVAIVGRPNVGSSTIFNRIVGERVSIYEDTPGVTRDRISYSGEMLTHDNTIDTGGIE 62
 DB 2 KPIIAIVGRPNVGSSTIFNLVGDKIAIVDLPGVTRDRILYRTDMSGSEFVIVDGGIE 61
 QY 63 ICGAPF-QTQIRAOAEIAIDEADVIIFMNVNREGILQSDDMAQIILKSKRPVLAIVNKV 121
 DB 62 PRNNDLMAIKIKQAEVAMEADVIILFVVDGKSLNPLDEIAYILIRKKRNPVILCNKI 121
 QY 122 DN-MENRTVDYDYFSLGFGEPYPISSHGILGLDLLDAVYSHRGEEDPDYDTRLSI 180
 DB 122 DNFEQDDYDYFSLGFGELVPISSHGKYNLGLMDLDIYVDIIGKMDPFEDEVLKLV 181
 QY 181 IGRPNVGSLLVNAIIGEDRVIVSNVAGTTRDAIDTEYSYDGYDVLIDTAGMRKGVY 240
 DB 182 IGRPNVGSLLVNAIIGEDRVIVSNVAGTTRDAIDTEYSYDGYDVLIDTAGMRKGVY 241
 QY 241 ESTEKSVLRALKATRSNNVIVDAEGCIIFQDKRVAGYAHQGNVIVVNMKDTVE 300
 DB 242 ESTEYYSVLRALKATRSNNVIVDAEGCIIFQDKRVAGYAHQGNVIVVNMKDTVE 301
 QY 301 -KDSKMKKFEEDVRRKFEQFDVDAQIAFVSAKERTLRTLFPIINASEMHHKRVSSSTL 359

DB 302 NKNNAIKKIKKEELVLAELPLSTAPIEFVSALTGORTNLEADRIYEYTRISTGL 361
 QY 360 NEVVTDIAISNPTPTDGRRLNVEYATQVAIEPPTFVFNVDLMHFYSKRYLENDIRA 419
 DB 362 NTILKDAVLMMNPPTTRGRVYIKINAYQVAVAPPKFVLCFNYPILHFSVARTENKRE 421
 QY 420 AFGEGTPIHI 430
 DB 422 AFGEGSPIMI 432
 RESULT 14
 ENG_AANSF STANDARD; PRT; 453 AA.
 ID ENG_AANSF
 AC O8Y2H7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR ALR0483.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CC NCBI_TaxId=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP003582; BAB72441.1; -
 DR PIR: AB1867; AB1867.
 DR HAMAP: MF_00195; -; 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP1_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR001806; Ras_transf.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR PRINTS: PR00326; GTP1_OBG.
 DR PRINTS: PR00449; RASTRNSFRMNG.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR TIGRFAMs: TIGR00231; small_GTP; 2.
 KM GTP-binding; Repeat; Complete proteome.
 FT NP-BIND 10 17 GTP 1 (POTENTIAL).
 FT NP-BIND 57 61 GTP 1 (POTENTIAL).
 FT NP-BIND 120 123 GTP 1 (POTENTIAL).
 FT NP-BIND 183 190 GTP 2 (POTENTIAL).
 FT NP-BIND 230 234 GTP 2 (POTENTIAL).
 FT NP-BIND 295 298 GTP 2 (POTENTIAL).
 SQ SEQUENCE 453 AA: 50731 MW: 23432F10C2E287D1 CRC64;
 Query Match 46.9%; Score 1037; DB 1; Length 453;
 Best Local Similarity 47.6%; Pred. No. 2e-56;
 Matches 210; Conservative 83; Mismatches 138; Indels 10; Gaps 5;
 QY 1 MKPIAIVGRPNVGSSTIFNRIVGERVSIYEDTPGVTRDRISYSGEMLTHDNTIDTGG 60

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:31 ; Search time 100 Seconds
(without alignments)
1125.109 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MTKPIVAIVGRPNVNGKSTIF.....IRAAFGPEGTPIHIIARRN 436

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2094	94.6	436	16 O8CP62	O8CP62 staphylococ
2	1614	72.9	436	16 O8EQA8	O8EQA8 oceanobacil
3	1523	68.8	436	16 O8E3T9	O8E3T9 streptococ
4	1523	68.8	436	16 O8DY73	O8DY73 streptococ
5	1476	66.7	436	2 O9RHV5	O9RHV5 streptococ
6	1472	66.5	436	16 O8DS90	O8DS90 streptococ
7	1342	60.6	445	2 O8GES8	O8GES8 heliobacill
8	1340.5	60.6	435	2 O8KH12	O8KH12 lactobacill
9	1028.5	46.5	449	16 O8DK11	O8DK11 synechococ
10	1022	46.2	444	16 O8EWH6	O8EWH6 mycoplasma
11	920	41.6	503	16 O8FF59	O8FF59 escherichia
12	898.5	40.6	496	16 O8LF58	O8LF58 vibrio vuln
13	874	39.5	537	10 O9LH52	O9LH52 arabidopsis
14	874	39.5	659	10 O9CF70	O9CF70 arabidopsis
15	859	38.8	487	16 O8EC36	O8EC36 shewanella
16	851.5	38.5	463	16 O8G6A8	O8G6A8 bifidobacte

17	843.5	38.1	552	16 O8FTK5	O8FTK5 corynebacte
18	838	37.9	588	16 O9A7R6	O9A7R6 caulobacter
19	796.5	36.0	483	16 O8G2E8	O8G2E8 brucella su
20	756.5	34.2	874	5 O8I5N5	O8I5N5 plasmodium
21	736	33.3	496	10 O9AW74	O9AW74 guillardia
22	678	30.6	489	16 O8F6K1	O8F6K1 leptospira
23	661	29.9	188	11 O9Z212	O9Z212 rattus norv
24	559.5	25.3	383	2 P94645	P94645 campylobact
25	524	23.7	642	10 O9AX00	O9AX00 oryza sativ
26	514.5	23.2	514	10 O9FLE0	O9FLE0 arabidopsis
27	512	23.1	347	2 O8VN97	O8VN97 heliobacte
28	512	23.1	347	2 O8VN36	O8VN36 heliobacte
29	511	23.1	347	2 O8VN38	O8VN38 heliobacte
30	455	20.6	281	2 O9LA73	O9LA73 thioacellu
31	435	19.7	456	16 O8DIY0	O8DIY0 walgieswort
32	433.5	19.6	208	2 P72548	P72548 synechococ
33	320.5	14.5	190	2 O9AE08	O9AE08 campylobact
34	320.5	14.5	190	2 O9Q016	O9Q016 campylobact
35	319.5	14.4	190	2 O9AE06	O9AE06 campylobact
36	318.5	14.4	190	2 O9AE09	O9AE09 campylobact
37	316.5	14.3	190	2 O9AE07	O9AE07 campylobact
38	266	12.0	479	16 O8DP28	O8DP28 streptococ
39	258	11.7	300	16 O8EPY0	O8EPY0 oceanobacil
40	243.5	11.0	455	16 O8DTT8	O8DTT8 streptococ
41	241.5	10.9	170	2 O9S3C9	O9S3C9 heliobacte
42	241.5	10.9	170	2 O9S3C8	O9S3C8 heliobacte
43	241.5	10.9	170	2 O9S3C7	O9S3C7 heliobacte
44	241.5	10.9	170	2 O8VNA0	O8VNA0 heliobacte
45	241.5	10.9	170	2 O9R3H9	O9R3H9 heliobacte

ALIGNMENTS

RESULT 1	ID	Query Match	Score	DB ID	Length
O8CP62	O8CP62	94.6%	2094	DB 16	436
AC	O8CP62	Best Local Similarity	93.1%	Pred. No. 2.9e-124	
DT	01-MAR-2003 (TRENBLrel. 23, Created)	Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;			
DT	01-MAR-2003 (TRENBLrel. 23, Last sequence update)				
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)				
DE	GTP binding protein.				
GN	SE1163.				
OS	Staphylococcus epidermidis.				
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
OX	NCBI_TaxID=1282;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 12228;				
RA	Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,				
RA	Chen Z., Wen Y.,				
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AE016747; AAC04760.1; -				
KW	Complete proteome.				
SEQ	SEQUENCE 436 AA; 49024 MW; 77C74B95641D3F22 CRC64;				

Query Match	Score	DB ID	Length
Best Local Similarity	93.1%	Pred. No. 2.9e-124	
Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;			
OY	1	MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIYEDPFGVTRDRYSSGEWLTDFDNITDGG	60
DB	1	MTKPIVAIVGRPNVNGKSTIFNRIVGERVSIYEDPFGVTRDRYSSGEWLTDFDNITDGG	60
OY	61	IEIGDAPFQTOIRAOAETAIADADVIFMVNVRGLTQSDENVAQILYKSKPPVLAANK	120
DB	61	IEIGDAPFQTOIRAOAETAIADADVIFMVNVRGLTQSDENVAQILYKSKPPVLAANK	120
OY	121	VDNMEMRDYVDFSLGCEPRYPSISGGLGLGLDLDVAVSHFGSEEDPDEDITRISI	180
DB	121	VDNMEMRDYVDFSLGCEPRYPSISGGLGLGLDLDVAVSHFGSEEDPDEDITRISI	180
OY	181	IGRPVNGSSLVNALIGEDRVIVSNVACTTDAIDTEYSDGQDVLIDTAGMRKKKRVY	240

```
DB 161 IGRPVNGSSLYNALGGERVYVSNVACTTRAIDITEVSYGQDYLIDTAGMRKKGY 240
OY 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
DB 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
OY 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
DB 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
OY 361 EYVTDIAISMNPPTDKGRRLNLFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
DB 361 EYVTDIAISMNPPTDKGRRLNLFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
OY 421 FGEGSTPIHIIARKRN 436
DB 421 FGEGSTPIHIIARKRN 436
```

RESULT 2

```
OQEOA8 PRELIMINARY: PRT; 436 AA.
ID OQEOA8:
AC OQEOA8:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE GTP binding protein.
GN OB1797.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_Taxid=182710.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004599; BAC13753.1; -.
KW Complete proteome.
SQ SEQUENCE 436 AA: 48998 MW: 18F52E40D9B0C9C2 CRC64;
```

Query Match 72.9%; Score 1614; DB 16; Length 436;

Best Local Similarity 69.4%; Pred. No. 5.8e-94;

Matches 302; Conservative 66; Mismatches 67; Indels 0; Gaps 0;

```
OY 1 MTKPIVAIVGRPNVSKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
DB 1 MKKSVAIVAGRPNVSKSTIFNRIYGERVSIYEDTPGVTRDRIRYANAEMLNHFENIIDGG 60
OY 61 IEIGDAPFOQIRAOAEIAIDADYIIFMVNVRREGLTOSDENVAOILYKSKRPVLAANK 120
DB 61 IEIGDAPFOQIRAOAEIAIDADYIIFMVNVRREGLTOSDENVAOILYKSKRPVLAANK 120
OY 121 VDNMEMRTDYYDFYSLGFGEPYPIGSHGLGDLDDAVSHFGEEEDPYDEDTIRLSI 180
DB 121 VDNMEMRTDYYDFYSLGFGEPYPIGSHGLGDLDDAVSHFGEEEDPYDEDTIRLSI 180
OY 181 IGRPVNGSSLYNALGGERVYVSNVACTTRAIDITEVSYGQDYLIDTAGMRKKGY 240
DB 181 IGRPVNGSSLYNALGGERVYVSNVACTTRAIDITEVSYGQDYLIDTAGMRKKGY 240
OY 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
DB 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
OY 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
DB 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
```

```
OY 361 EYVTDIAISMNPPTDKGRRLNLFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
DB 361 EYVTDIAISMNPPTDKGRRLNLFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
OY 421 FGEGSTPIHIIARKRN 435
DB 421 FGEGSTPIHIIARKRN 435
```

RESULT 3

```
OQE379 PRELIMINARY: PRT; 436 AA.
ID OQE379:
AC OQE379:
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN GBS1667.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_Taxid=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangoul L.,
RA Msadek T., Zouine M., Couve E., Lallouf L., Poyart C., Tiliou-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766852; CAD47326.1; -.
KW Sagalists; gbs1667; -.
SQ Hypothetical protein; Complete proteome.
SQ SEQUENCE 436 AA: 48880 MW: E7E33B4D0F3F61G*GRC64;
```

Query Match 68.8%; Score 1523; DB 16; Length 436;

Best Local Similarity 66.5%; Pred. No. 3.2e-88;

Matches 290; Conservative 70; Mismatches 74; Indels 2; Gaps 2;

```
OY 1 MTKPIVAIVGRPNVSKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
DB 1 MTKPIVAIVGRPNVSKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
OY 61 IEIGDAPFOQIRAOAEIAIDADYIIFMVNVRREGLTOSDENVAOILYKSKRPVLAANK 120
DB 61 IEIGDAPFOQIRAOAEIAIDADYIIFMVNVRREGLTOSDENVAOILYKSKRPVLAANK 120
OY 121 VDNMEMRTDYYDFYSLGFGEPYPIGSHGLGDLDDAVSHFGEEEDPYDEDTIRLSI 180
DB 121 VDNMEMRTDYYDFYSLGFGEPYPIGSHGLGDLDDAVSHFGEEEDPYDEDTIRLSI 180
OY 181 IGRPVNGSSLYNALGGERVYVSNVACTTRAIDITEVSYGQDYLIDTAGMRKKGY 240
DB 181 IGRPVNGSSLYNALGGERVYVSNVACTTRAIDITEVSYGQDYLIDTAGMRKKGY 240
OY 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
DB 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
OY 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
DB 301 KDSKTMKKFEDEVKREFOFLDYAOIAFVSAREKTRTLRFLPYINASENHKKRVOSSTLN 360
OY 420 AFGEGSTPIHIIARKRN 435
DB 420 AFGEGSTPIHIIARKRN 435
```


RESULT 4	ID	PRELIMINARY:	PRJ:	436 AA.
08DY73	08DY73			
AC	08DY73;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Phosphoglycerate dehydrogenase-related protein.			
GN	SAG1620			
OS	Streptococcus agalactiae (serotype V).			
OC	Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=216466;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2603 V/R / Serotype V;			
RX	MEDLINE=22222988; PubMed=12200547;			
RA	Tetelin H., Maignanl V., Ciesliewicz M.J., Eisen J.A., Peterson S.,			
RA	Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,			
RA	Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,			
RA	Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,			
RA	Radune D., Pedrovica N.B., Scanlan E., Khouli H., Mulligan S.,			
RA	Carty H.A., Cline R.T., Van Aken S.E., Gill J., Searelli M., Mora M.,			
RA	Iacobini E.T., Brettoni C., Gelli G., Mariani M., Vegni F., Malone D.,			
RA	Rinaudo D., Kappoli R., Telford J.L., Kasper D.L., Grandi G.,			
RA	Fraser C.M.;			
RT	"Complete genome sequence and comparative genomic analysis of an			
RT	emerging human pathogen, serotype V <i>Streptococcus agalactiae</i> ."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).			
DR	EMBL: AE014265; AAN00484.1; -			
DR	TIGR: SAG1620; -			
KW	Complete proteome.			
SQ	SEQUENCE 436 AA; 48981 MW; E2C064B3F04B8644 CRC64;			

Query Match	68.8%	Score 1523	DB 16	Length 436
Best Local Similarity	66.3%	Pred. No. 3,2e+88		
Matches 289	Conservative 72	Mismatches 75	Indels 2	Gaps 2
QY	1	MTKPIVAIGRPNGKSTIFNRIVIGERSVIEDTPGVTRDRIRYSSEWMLTHDFNIIDTGG	60	
Db	1	MVLPTVAIVGPRNGKSTPLFNRIAGERISIVEDGVGVRDRIRYTTGEMLNKRFSLIDTGG	60	
QY	61	IEIDAPRQGTIRKQAEIIAIDEADVIIIFMVAVRSLTQSDMEVNAQILTKSKKPPVLAANK	120	
Db	61	IDVDAPRMEGIKHQADIAAMEADVIYFVVGSKESGVTDADEVSRILKTKNPPVLAANK	120	
QY	121	VDNNEMRDVIYDFSLGFGEPYPISGSHGLIGLDLDAVSHFGEEEDPDDEDTIRLSI	180	
Db	121	VDNEMNRDVIYDFSLGIGDPPPLSSVHGIGTGDLDAIVENLPEEENE-NPDLIRSL	179	
QY	181	IGRPNGKSSLVNAIIGEDRAVIVSNVACTRDADITEX-SYDGODVYLIDTAGMRKKGV	239	
Db	180	IGRPNGKSSLINILGEDRAVIVSVAAGTDRADITDNEVDSOGGEYTMIDTAGMRKSGKV	239	
QY	240	YESPEKYSVLKALAIERSNVVLVYIDAEGLIEDDKRVACAGHOGKAAVYVVKMPTV	299	
Db	240	YENEEKISYMSKMRIDRSVDVLAIVTNAEEGRREYDKRIAGAHHTGGLIIIVKMPDTI	299	
QY	300	EKDSKTMKFFDEYAKREFQFLDVAQIAVASAKERTRLTLPYIMEASENHKRVQSSLT	359	
Db	300	EKDNHTYSQWADLRDNQGFSLYAPRIIFVSAETKRLKRLKLPDMKIRIESQNRKIRPSAVL	359	
QY	360	NEVYTDIAISNPTPDCKGRRLNIFYAIOVAIEPPTFVVVFVNDVELIMHSYKRYLENOIRA	419	
Db	360	NDVIMDAIAINPTPDCKGRRLKIFYAIOVAIVKPPFVVVFVNEEELIMHSYLFLENOIRE	419	
QY	420	AFGEGFPPIHIIARRK 435		
Db	420	AFVEEGPIINLIARRK 435		

```

09RHV5
ID 09RHV5          PRELIMINARY;          PRT;          436 AA.
AC 09RHV5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Phosphoglycerate dehydrogenase.
GN PCDA.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MT8148;
RA Kawabata S., Terao Y., Hamada S.;
RT "Molecular cloning, sequence and characterization of a novel
RL Oral Microbiol. Immunol. 15:58-62(2001).
DR EMBL; AB016077; BAA8823.1;
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OBG.
DR TIGRFAMs; TIGR00650; MG442. 2.
DR TIGRFAMs; TIGR00231; small_GTP. 2.
SQ SEQUENCE 436 AA; 48601 MW; EC52CC8650DE090 CRC64;

Query Match          66.7%; Score 1476; DB 2; Length 436;
Best Local Similarity 62.8%; Pred. No. 3e-85;
Matches 274; Conservative 83; Mismatches 77; Indels 2; Gaps 2

```

Query Match	66.7%:	Score 1476:	DB 2:	Length 436:
Best Local Similarity	62.8%:	Pred. No. 3e-85:		
Matches 274:	Conservative 83:	Mismatches 77:	Indels 2:	Gaps 2:
QY	1	MTKEIVAIVGHPNNGKSTIFNRIIGERVSIYEDTPGVTRDRIRYSSEWMLTHFDNIIDTG	60	
Db	1	MALEPVAIVGHPNNGKSAFLPNRIAGERISIVEDVGVTRDRIRYTKAEWLNRFQSIITDGG	60	
QY	61	IEIGDAPQOTQIRKQAEIAIDEDAVYIIFMVAVREGILQSDENVAOILTKSKKPPVLAANK	120	
Db	61	IDVDAPPEMEQIKHQADIAMTEADYIVVSAKEBITADEYVAKILYRTHKPPVLAANK	120	
QY	121	VDNNEMRDVPDYFSLGEGPEYPISGSHGLGLDLDAVSHFGEEDPYDEDTIRLSI	180	
Db	121	VDNEMRSARIDYFALGLGDPYPPYSSAIGICGVDLVAIVDNLPAQAGE-SSDLTKRSL	179	
QY	181	IGRPVNGKSSLVNAILGEDRIVSVNAGTTRDAIDTEX-YDGQDYLDITAGMKKGV	239	
Db	180	IGRPVNGKSSLINAILGEDRIVASVAGTTRDAIDTFTTDEGQEFMTIDFAGMKRSGV	239	
QY	240	YESPEKSVLALKAIESNNVLYVYDAEOGIIEDKRVAGAAHOGKAAVYVYVKKMPDV	299	
Db	240	YENEEKSVYMMARRIDISDVIYMLWLNABEGIREYDKRIAGAHBAGIGVYVYVKKMDYI	299	
QY	300	EKDSYMKFEDEYAKKEFOPLDYAOIAVSAKERTRLTLPYIMEASENHKKRVQOSTL	359	
Db	300	KKDRRTVAQWETDIRDNQYIPYAPRIYVSVAVTKRRLKLPDVIKQVQSQMTFLPSSVL	359	
QY	360	NEVVYTDASINPTPLDKGRRLNVFYATOVAIEPPFFVYFVNDVDELMHFSYKYLENQIRA	419	
Db	360	NDVVMDAAVAIPTPLDKGRRLKIFATOVSVKPPFFVIFVNEDEELMHFSYLFLENOIRQ	419	
QY	420	AFEGEGPPIHIIARKR	435	
Db	420	AFVEEGPIRLIARKR	435	

RESULT 6

Q8DS90 PRELIMINARY: PRT: 436 AA.

AC Q8DS90.

DT 01-MAR-2003 (TTEMBLrel. 23, Created)

DT 01-MAR-2003 (TTEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE Phosphoglycerate dehydrogenase.
GN PGDA OR SMU.1920.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
RN NCBI_TaxID=1309;
RP
RC SEQUENCE FROM N.A.
RX STRAIN-UAI59 / ATCC 700610 / Serotype C;
RA MEDLINE=22295063; PubMed=12397186;
RA Aylde D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL: AE015016; MAM59531.1; -.
KW Complete proteome.
SQ SEQUENCE 436 AA; 48585 MW; 24D8428A91C2A097 CRC64;

Query Match 66.5%; Score 1472; DB 16; Length 436;
Best Local Similarity 62.6%; Pred. No. 5.3e-85;
Matches 273; Conservative 84; Mismatches 77; Indels 2; Gaps 2;

QY 1 MTKPIVAIYGRPNVSKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTDFNITDGG 60
DB 1 MALPVAIYGRPNVSKSALFNRIAGERISIVEDGVTRDRIRYTAEMLRQFSIIDTGG 60
QY 61 IEIGAPFQTOIRAOAETIADADVIIFVNVNREGITQSDENVAQILKSKRPVLAIVK 120
DB 61 IDDVAPFMEQIKHQADIAIMTADVIIFVNSAKGKITDADEVAKILYTHKRPVILAIVK 120
QY 121 VDNMEARTDYDFYSLGEGEPYPIGSHGLGDLDAVAVSHFGEEDPYDEDTIRLSI 180
DB 121 VDNPEKRSIYDFYALGLGDPYFVSASRGITGVDVLAIVDNLPEADQE-SSDIKESL 179
QY 181 IGRPNVGSLLVNAIIGEDRIVSNVAGTTRDAIDTEYS-YDGQDVLIDTAGMRKKGY 239
DB 180 IGRPNVGSLLVNAIIGEDRIVSNVAGTTRDAIDTEYS-YDGQDVLIDTAGMRKKGY 239
QY 240 YESTKYSYLRALKAIERSNVVLYVIDAEOGIIEDDKRAGVAGHAGKAVVIVNKKMDV 299
DB 240 YENTKYSYLRALKAIERSNVVLYVIDAEOGIIEDDKRAGVAGHAGKAVVIVNKKMDV 299
QY 300 EKDSKTKMKFEDEVKREFQFLDYAOIAFVSAKERTRLTLPYINEASENHKKRYOSTL 359
DB 300 KKDNTFVAGMEADIDNDQIYIPAYIVFSAVTKORLHKLPVYIKVQSOSQMTRIPSAVL 359
QY 360 NEVTDAISMNPPTDKGRRLNVFAOVAIEPPTFVFNVDVLMHFSYKRYLENQIRA 419
DB 360 NNVVMDAVAINPTPDCKGRKRIKFATQVSVKPPFVIVNEEELMHFSYLRLENQIRQ 419
QY 420 ARGEFTPIHIIARRK 435
DB 420 AFVEGCTPIRLIARRK 435

RESULT 7
08GES8 PRELIMINARY: PRT: 445 AA.
AC 08GES8; 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE GTP-binding protein (Fragment).
OS Helicobacillus mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;
OC Helicobacillus.
OX NCBI_TaxID=28064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22337798; PubMed=12446909;

RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
RA Blankenship R.E.;
RT "Whole-genome analysis of photosynthetic prokaryotes."
RL Science 298:1616-1620(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
RA Gerdes S., Kyriides N., Overbeek R.;
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY142760; MAM87364.1; -.
FT NON_TER 445 445
SQ SEQUENCE 445 AA; 49885 MW; 1D6A692E387AD61C CRC64;

Query Match 60.6%; Score 1342; DB 2; Length 445;
Best Local Similarity 56.8%; Pred. No. 8.8e-77;
Matches 250; Conservative 83; Mismatches 99; Indels 8; Gaps 3;

QY 1 MTKPIVAIYGRPNVSKSTIFNRIVGERSIVEDTGVTRDRIRYSSGEMLTDFNITDGG 60
DB 4 MAKPIVAVVGRPNVSKSTIFNRLTGGRVAIVEDRGVTRDRIRYRNAKWLNRFTLVDTGG 63
QY 61 IEIG--DAPFQTOIRAOAETIADADVIIFVNVNREGITQSDENVAQILKSKRPVLAIV 118
DB 64 IERGADGNPFSEVYIKQAEATAIADADVIIFVNDGKAGITADETYAAVLRRTKRYFLVY 123
QY 119 NRVNMEKRTDYDFYSLGEGEPYPIGSHGLGDLDAVAVSHF--GEEDPYDEDT 175
DB 124 NKIEDPSQDKYEFEPFALGLGDPYFVSASRGITGVDVLAIVDNLPEADQE--DT 180
QY 176 IRLSIIIGRPVSKSSLVNAIIGEDRIVSNVAGTTRDAIDTEYSIDGQDVLIDTAGMRK 235
DB 181 IKIAVIGKPNVSKSSLVNAIIGEDRIVSNVAGTTRDAIDTEYSIDGQDVLIDTAGMRK 240
QY 236 KGVYESTKYSYLRALKAIERSNVVLYVIDAEOGIIEDDKRAGVAGHAGKAVVIVNKK 295
DB 241 KKRIDESVRYSKMSLRVDSVDVLYMIDASQGTEDDKRIAGAHAGACVLYLVNK 300
QY 296 WDTVEKSKTKMKFEDEVKREFQFLDYAOIAFVSAKERTRLTLPYINEASENHKKRYO 355
DB 301 WDLVPRDXTMNRFRQDVVSEGLFNAPYIVASALTGQRLPIELVDFVFEQANRRIP 360
QY 356 SSTLNENVYDAISMNPPTDKGRRLNVFAOVAIEPPTFVFNVDVLMHFSYKRYLEN 415
DB 361 TSVINELMDIIRVPAPSDRGRKILVYTOTSVKPPFVFNVDVLMHFSYKRYLEN 420
QY 416 QIRAAFGFTPIHIIARRK 435
DB 421 RREFTGFGCTPVRLIARR 440

RESULT 8
08KH12 PRELIMINARY: PRT: 435 AA.
AC 08KH12; 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE Conserved hypothetical GTP-binding protein.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 11842;
RA Serrro P., Deryn R., Ehrlich S.D., Maguin E.;
RT "Lactobacillus delbrueckii spp. bulgaricus hba region."
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY094626; MAM22484.1; -.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.

Query Match	60.6%	Score 1340.5	DB 2	Length 435
Best Local Similarity	59.1%	Pred. No. 1.1e-76		
Matches 257	Conservative 68	Mismatches 109	Indels 1	Gaps 1
QY	1	MTKPIVALVIGRPNNGSKSTFNRIYGERVSIYEDPFGVTRDRRIYSGEVLITDFNIIIDGG	60	
Db	1	MLPLVIALVGPNGSKSTLFNRIILINERVAIVEDRGVTRDRYARASWGHQFSLIDTGG	60	
QY	61	IEIGDAFQNTQIRAAQAEIAIDEADVIIFPMVVRREGITQSDENVAOILYKSKKPVYLAANK	120	
Db	61	ITWEDSTIDEIRRAQAEIAIEADVIYVLAASQVSTSLDERIALILYKRAKPVYLAANK	120	
QY	121	VDNMKARTVDYDFYSLGFGPEPPIYSGSHGLGLDLDVAVSHFGSEEDPDYDEDTIRLSI	180	
Db	121	ADNEQGRFDIYDFSLGGLDPIPVSGSHGTGLDLDLVYKFNFSDEAKT-EEGVISFSV	179	
QY	181	IGRPNVGSLSLVNAILGCDRIYVSNVATTTDADITDESYGQDVLVLDTAGMRKKGVY	240	
Db	180	IGRPNVGSLSLVNRIIGERVIYVANEESTTDDAIDTPVKQGTGRVVDTAGIRRRGVY	239	
QY	241	ESTERYSYLRLAKAIEERSNVVLVVIDAEOGIIEDQRAGYAEHOGKAVIYVNMKRVYE	300	
Db	240	EKTERYSVMRAMSAMERSDVALIVLDASTGIREQDKHAGYAHENGLOMIITAVNMKDLPK	299	
QY	301	KDSYTKMKKEDEVRKEQFQFDYAOIAYVSAKERTRLRFLPYINPASENHKKRVOSTLN	360	
Db	300	KDSSSGKDFEAVIYREEFSYLDAPYFAISAKTGKNIDOLPKMKVEYENKKNORIOSSVLN	359	
QY	361	EVVVIDAISMNPTPDMDKGRRLWFAVATOAIAIEPPFPVNVVNDYELMHFSKKRYLENOIRA	420	
Db	360	DLLLEASGLVLPAPMYKGRRLVYVYTVQVKNPFPVFCNDPELHMFSTORFLINOLREN	419	
QY	421	FFEGSTPIHIIARRK 435		
Db	420	FDFTGTPIKILPRKR 434		
RESULT 9				
Q8DKI1	PRELIMINARY:	PRT:	449 AA.	
AC	Q8DKI1:			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	GRP-binding protein.			
GN	TLR0878.			
OS	Synechococcus elongatus (Thermosynechococcus elongatus).			
OC	Bacteria: Cyanobacteria: Chroococcales; Synechococcus.			
OX	NCBI_TaxID=32046;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RC	STRAIN-BP-1;			
RX	MEDLINE=22225144; PubMed=12240834;			
RA	Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,			
RA	Watanabe A., Iriyuchi M., Sawashima K., Kimura T., Kishida Y.,			
RA	Kiyokawa C., Kohara M., Matsuno M., Matsuno A., Nakazaki N.,			
RA	Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;			
RT	"Complete genome structure of the thermophilic cyanobacterium			
RL	Thermosynechococcus elongatus BP-1.";			
RL	DNA Res. 9:123-130(2002).			
DR	EMBL; AP005371; BAC08430.1; -			
KM	Complete proteome.			
SO	SEQUENCE 449 AA; 50049 MW; 4109222725A655D5 CRC64;			

Query Match	46.5%	Score 1028.5	DB 16	Length 449
Best Local Similarity	47.6%	Pred. No. 5.5e-57		
Matches 209	Conservative	87	Mismatches 136	Indels 5
QY	1	MTKEPIVAIVGPNNGKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEWLTHDNFIIDTGG	60	
DB	1	MALEIIVAVGPNKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEWLTHDNFIIDTGG	60	
QY	61	IEI-GDAEPQTQIRAAQETIADEADVITFMVNVREGLTQSDENVAOILYKSKRPVVLAVN	119	
DB	61	LVFDDSDSDFLLIRQAQELALQEAFTAAIFVVDGQAGPTALDEYIAAMLROLSLPVLAVN	120	
QY	120	KVDNMKM-RITVDYDFYSIGFGEPRYITSSHGIGLGDLLDAVYSHNGEEDREDDTJRL	178	
DB	121	KCESRQMQOYAAEFWSJGLGEPRYISSIHSGSTELLDQLTTPYAGSTLP-EAREIYV	179	
QY	179	SIISRPNGKSSILVAALIGEDRVVYSNVAGTTTADITDEYSYDGDVYLIDTAGMRKKK	238	
DB	180	AIAGRPNGKSSILNALIGSDRAIVSPISGTTROADIDVYIENGQYQYRITDPAIGIRKTH	239	
QY	239	VEESTEKYSVLRAIKAIERSNVVLVDAEGQIEBDRKVAGYAHQSKAVVIVVKKWDT	298	
DB	240	VAYGGEHMSVHRAFAKAIHRSDVLLVLDLAEITBODORLAGHIDQGRACVLIYKKMDA	299	
QY	299	V-EKDSKTMYKFEDEVEKKEFQFLDYQAQIAFSAKERTYRLTFLPTINASESHKKRVQSS	357	
DB	300	VLDKDYITAYINAYRRLYORLHFLHFWADLAFVSAHNGORLEKFAAVDAVDEHRRVYTA	359	
QY	358	TLNEVVTPTAISMN-PTTPDKGRFLVWFATQVAIEPPTFVVNVNDVLMHFSYKRYLHQ	416	
DB	360	VVNDVLDQALMHHTPRPATRQGRQGIYATQVATQPTPTAFITVNDAKLFKENVKRYTIESQ	419	
QY	417	IRAAFGEGTPTPIHIAKKR 435		
DB	420	IRQQLGFRGTPRIILMRSK 438		
RESULT 10				
ID	Q8EWH6	PRELIMINARY	PRT	444 AA.
AC	Q8EWH6			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
GN	Predicted ATP/GTP-binding protein.			
OS	Mycoplasma penetrans.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_taxID=28227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HF-2;			
RX	MEDLINE=22354719; PubMed=12466555;			
RA	Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,			
RA	Yoshino C., Horino A., Shida T., Sasaki T., Hattori M.;			
RT	"The complete genomic sequence of Mycoplasma penetrans, an			
RL	intracellular bacterial pathogen in humans."			
DR	Nucleic Acids Res. 30:5293-5300(2002).			
KW	EMBL; AP0041170; BAC44020.1; -			
SO	Complete proteome.			
Query Match	46.2%	Score 1022	DB 16	Length 444
Best Local Similarity	47.8%	Pred. No. 1.4e-56		
Matches 207	Conservative	80	Mismatches 140	Indels 6
QY	6	VAIVGRPNVGSSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEWLTHDNFIIDTGGTIED	65	
DB	4	VAIVGRPNVGSSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEWLTHDNFIIDTGGTIED	63	
QY	66	APFQOIAQAQETIADEADVITFMVNVREGLTQSDENVAOILYKSK-RKVVLAVNVDM	124	
DB	64	EDFKQIQLQVNVVATTEADVIVLVSYEGKINNDHYKALIKKHNKRIITLLAVNKESTR	123	

[illegible]

```

RESULT 11
08FF59 ID 08FF59 PRELIMINARY; PRT; 503 AA.
AC 08FF59;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Probable GTP-binding protein enga.
CN C3033.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL: AE016764; NAMB1483.1; -.
KW Complete proteome.
SQ SEQUENCE 503 AA; 56573 MW; 711A9BD7863280A CRC64;

```

Query Match	41.6%;	Score 920;	DB 16;	length 503;
Best Local Similarity	42.0%;	Pred. No. 4.6e-50;		
Matches 197;	Conservative 91;	Mismatches 143;	Indels 38;	Gaps 4;

```
OY      4 IVALVAGPNNKSTIPFRNIGRESIVEDTFCGVNRDRIRYSSEGMELTHDNIIDGCIET 63
       | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     16 PVALVGPRPNKGSTLEFNLTTRIALVADEPGTLTRDKRYGRAELIEGRFICDTGTGDIG 75

OY      64 GDAPFOQIIRAQAETAIDEADVIIFMVNVNREGLTOSDENAVAOILYKKRPVLVAVKVDN 123
       | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     76 TEDGETHMAEGSLAIEBADVYLPMVOARALMPADELAIAHLRBRERPTFLVAKKTDO 135

OY      124 MEMRDVDFYSLGFGEPPISGSNGLGIDLDAVV----- 160
       :: | |||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     136 LDPDDAAVDFAALGGELYPIASHGRCVLSLEHHLLPWMDLIAQEVEDEDAERYWAOF 195

OY      161 ---SHFGEEEDPYEDFT--IRLSITGRPVNKSSLYVNAILGEDRVYSVMAGTTBDALD 215
       | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    196 EAENGEEEBEEDFPPOSUPLKILAIVAGRNVKSTJTNKLILEERVYVDMPTKDOSTY 255
```

```
OY      216 TETSYGQDYLVDLDFGAMKKKGVSTPTKYSVLALAKAIERSNNVLVVIDAEGIIED 275
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      256 IPERDSREVLLDTAGVRCKGKITDAVEKFSSVITFDLAIDBANVMVLVIDAREGISDD 315
OY      276 KRVAGYAHBOGKAIVLVIVNKKMDPTVERKDOSTYMKKFEDEVKRKEFOFLDYAOIAFYASAKERTR 335
          :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      316 LSLGLFTLNSGRSLVTVNKKMDELSEGEVQYKELTLDR--LGTFIDFARHVFIALHGSG 373
OY      336 LRTLPFIYNASENHKKRVQOSSTFLNEVTDALISNNPFTDKGRLNLNFYATQVAIEPPTF 395
          ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY      374 VGNLFESVREAYDSSSTRVGTSMLTPIIMTMAVEDHQPLVNGRRVKLEYYAHAGGYNPPIV 433
Db      396 VFVFNDELMHFYSKYRYLENOIRAAFGECGRPIHI-----IARRKN 436
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
OY      434 VIHNQOVKLDLPDSTKRYTLNNFRKSLDWVGSRITQLQEGENPYANKKN 482
```

RESULT	12
ID	08DF02
AC	08DF02
DT	PRELIMINARY;
DR	PRT; 496 AA.
DE	01-MAR-2003 (TRENBLREL. 23, Created)
DE	01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DE	01-MAR-2003 (TRENBLREL. 23, Last annotation update)
GN	GTP binding protein.
OS	VV10423.
OC	Vibrio vulnificus.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=672;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-CMCP6;
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.Y., Moon Y.H., Jeong H.,
RA	Choy H.E.;
RT	"Complete genome sequence of Vibrio vulnificus CMCP6.";
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AE016798; AAC08946.1; -
KW	Complete proteome.
SO	SEQUENCE 496 AA; 55469 MW; 271966BA312B7923 CRC64;

Query Match	40.6%	Score	898.5	DB	16	Length	496
Best Local Similarity	42.4%	Pred	No. 1e-48				
Matches	195	Conservative	84	Mismatches	146	Indels	35
				Gaps			4
QY	4	PVALVGRPNVNGSKSTFIENRIVGERVSIEDTPGVTRDRIRSYSGEWLTHDFNIITDGGIEI	63				
DB	3	PVALVGRPNVNGSKSTFLFNRLTRSRDALVADPFGLTRDRKRGCAKVGHEIDFVIDTGGIDG	62				
QY	64	GDAPFQIQIRQAELAIDEADVIITPMVAVRELTQSDENVNOILYKSKKPPVLAANKVDN	123				
DB	63	SEEGVETMAQDSLAIREADVILFPMVGRAGLTPTSDELAHLKIEKATMLVYNNKVDG	122				
QY	124	MEMRTDVPYFSLGGEPEYPIGSHGLDGLDVAVSHF-----GEEE--DPY	171				
DB	123	IDADASADPFLQGLVDENYQIAAAGRGVLTALIEALDRPFNNLLSANNEGIEDLITDME	182				
QY	172	DEDT-----IRLSITGRPNVNGKSSLYNALILGEDRVIVSVAGTT	210				
DB	183	DEDAEQGEYSEDAEESLKRLODQPIKALIIIGRPVNGKSTFLNRLIGERVVVYDMPGTT	242				
QY	211	RDAIDTEKSYSGODVILDTAGMRKKGVYESTEKYSVLRLAKAIERSNVVVLVIDADRG	270				
DB	243	RDSYIIPERDQREVLIDTAGVRRKGVVHEVEFVSVAKTLKAEDAAVALLVIDAREN	302				
QY	271	IIEDKRVAGYAHEOGAKVAVIVNNKMDVVEKDSKTMKKEFEDEVKREFOLDYAQATFASA	330				
DB	303	ISDQDLSLGFLNAGRSIVLAVNNKMDGL--DNEVKENMKELDRGLRGPVDEARLHFISA	360				
QY	331	KERRRLRLTFPIYNASENHHKKRQVSTFLNEVYVDAISMNPPFDKGRRLNVFYATQVAL	390				
DB	361	LHGVGVLGHFSVGEYASATTRVGTSVLTRIMKAAATDHDQPPMVRGRVRYKJYKTHAGY	420				


```

QY 62 -----EIGDAFPOTQIRAOAETIADEADVIIFMVNVRG 95
DB 217 KSPSGVMEELNVSTTIGMEGIPLSRREAIAIRHPSMIEKQATAVDESAVIIEFVDDQAG 276
QY 96 LTQSDENVAAQIL--YKSKRRVVLAVNKNVDN-MEMRTDVIYDFYSLGFEPPPIGSHGL 152
DB 277 PSQADVEIADWLKRYYSKTYIIILAVNKCSPRGLMQASEFWSLGF--TPPIALSAGTGT 335
QY 153 GDLLDAVVS-----HFGEEEDPDYEDDTIRLSITGRPNVGSLSLVNAILGEDRVIV 203
DB 336 GELLDLVCSGLIKLEIMENEIEEEENY---IPALITGRPNVGSLSILNALVEDRTIV 392
QY 204 SNAAGTTRDAIDTEYS-YDGQDYVLIDTAGMRKKGVES---TEKYSVLRAKAIERSN 259
DB 393 SPVSGTTRDAIDAEFTGPDEKFRLLIDTAGIRKSSVASSGSTTEAASVNRAPRAIRSD 452
QY 260 VLVIVLDAEAGIIEODKRVAGYAHGOKAVIIVNKKDVI-EKDSKMKKFEDEVKREPO 318
DB 453 VVALVIEAMACITEODUKIAERIEREGKGLVYVNNKMDTIPNKKQETAAHVEDVREKLR 512
QY 319 FLDYAQIAFVSAKERTRLTLFPYINEASENHHKRVOSTLNEVYTDALISM-NPTPYDKG 377
DB 513 SLKMAPIVSTALITGHSVDNIVVAATVQKERSRLSTALLNQVIRAAVAFKSPRRRG 572
QY 378 RRLNVFATQVAIEPPFFVYVNDVELMHFSYKRYLENOIRAAFGFEGTPIHIARR 435
DB 573 KRGVYVYCTQAAIRPPTFEVFNDAKLFSDTYRRYMEKQIRTAGFAGTPIRLMRSR 630

```

RESULT 15

```

ID 08EC36 PRELIMINARY; PRT; 487 AA.
AC 08EC36;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE GTP-binding protein EngA.
GN ENGA OR SO3308.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxId=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Ufflerback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.,
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL: AE015769; AAN56306.1; -.
DR TIGR: SO3308; -.
KW Complete proteome.
SQ SEQUENCE 487 AA; 54531 MW; B75F86CA1BF9C20 CRC64;

```

Query Match 38.8%; Score 859; DB 16; Length 487;
 Best Local Similarity 39.9%; Pred. No. 3,1e-46;
 Matches 180; Conservative 98; Mismatches 147; Indels 26; Gaps 4;

```

QY 4 PIVAIAGRPVNGKSTINRIVGERVSTVEDTGVTRDRIRYSSGEMLTGHDENIIDTGIEI 63
DB 3 PVALVIGRPVNGKSTLNRIRLRLTRDALVADFPGLTRDRKYGRAFLSGYEIVVDGIDG 62
QY 64 GDAFQOIRAOAETIADEADVIIFMVNVRGGLTQSDENVAAQILYKSKRRVVLAVNKNVDN 123

```

```

DB 63 TEEGIEFKMAEOSTIAIEADVYLFMTDARAGLTAADLSIAOHLRSROKTFPVVANKIDG 122
QY 124 MEMRTDVIYDFYSLGFEPPPIGSHGLGLDLLDAVVSHP-----GEEB---EDPY 171
DB 123 IDADSACAEFWISGLGEVYQMAAAQGRGVNMIETVALTPYAAMGIERQGEEDERQY 182
QY 172 DED-----TIRLSITGRPNVGSLSLVNAILGEDRVIVSNVAGTTRDAIDTEYS 219
DB 183 TEEEAEEAQRLQDLPIKLAIIKGPVNGKSTLNRILGEEBRYVYDEPGTTRDSIYI 242
QY 220 YDGQDYVLIDTAGMRKKGVESYESTKYSVLRALKATERSNVVLVIDAEGIIEODKRV 279
DB 243 RDGREYIIDTAGYRRSKVHEVIEKFSVTKTKAAVEDANVLIIDAREGVAEQDLGL 302
QY 280 GYAHGOKAVIIVNKKDVIYKDSKTKKFEDEVKREFOPLDYAQIAFVSAKERTRL 339
DB 303 GFALNAGRALVIAVNNKWDGIDQGIK--DRVKSELDRLGFIIDFARIHFISALHGTG 360
QY 340 FPYINEASENHHKRVOSTLNEVYTDALISMNPTPYDKGRLLNFYATQVAIEPPTFV 399
DB 361 FESIEEAYDSATRRVSTSMLTRIMQMSQDDHPPVLVNGRRVKKLVYAHAGVNPPIV 420
QY 400 NDVELMHFSYKRYLENOIRAAFGFEGTPIHI 430
DB 421 NOVSKLPDSYKRYTMNFRSLKAVYGTPIQL 451

```

Search completed: September 10, 2003, 00:35:04
 Job time : 105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:16:39 ; Search time 129.915 Seconds
(without alignments)
8041.284 Million cell updates/sec

Title: us-09-815-242-1463

Perfect score: 387
Sequence: 1 gatcttcctctcttcacca.....ttcacctgaagaataaac 387

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	387	100.0	387	AAS48886
2	387	100.0	1311	AAT86461
3	387	100.0	1311	AAS54997
4	383.8	99.2	1305	AAS51646
5	383.8	99.2	1311	AAS54865
6	327	84.5	3621	AAV74669
7	295	76.2	372	AAS50706
8	262.2	67.8	1332	ABN50883

9	262.2	67.8	3269	22	AAS4708
10	234	60.5	298	23	AAS50205
11	234	60.5	298	23	AAS50723
12	205.8	53.2	319630	24	ABO67194
13	205.8	53.2	2944528	24	ABO03041
14	205.8	53.2	3011208	24	ABO69245
15	174.4	45.1	1311	24	ABK75008
16	162.6	42.0	1308	24	ABN68458
17	162.6	42.0	2365589	24	ABA90521
18	160.4	41.4	1308	24	ABN68457
19	160.4	41.4	2155561	24	ABN71527
20	153	39.5	1311	21	AAS4516
21	151.4	39.1	1308	25	ABX07474
22	151.4	39.1	1311	21	AAZ91826
23	151.4	39.1	5066	19	AAZ5212
24	151.4	39.1	2162598	25	AB56454
25	146	37.7	246	23	AAS49268
26	101	26.1	960	22	AAS5683
27	89.4	23.1	1512	23	AAS52517
28	89.4	23.1	11574	22	AAS46244
29	86	22.2	157	23	AAS50894
30	84.4	21.8	1512	22	AAF94379
31	84.4	21.8	1512	24	ABK64943
32	84.4	21.8	1515	23	AAS53235
33	83.8	21.7	1473	23	AAS55968
34	80.6	20.8	640681	24	ABA92787
35	80.2	20.7	1830121	17	AAT42063
36	73.2	18.9	1370	19	AAZ96377
37	73	18.9	33140	22	AAF28536
38	69.8	18.0	1146	25	ACA00899
39	69.8	18.0	1557	22	AAH65442
40	69.8	18.0	349980	22	AAH68528
41	69.8	18.0	349980	22	AAH68529
42	64	16.5	25360	22	AAF88314
43	64	16.5	29736	22	AAF88317
44	62	16.0	580073	18	AAT58840
45	61.4	15.9	78845	21	AAA81463

ALIGNMENTS

RESULT 1
AAS48886
ID AAS48886 standard; DNA; 387 BP.
AC AAS48886;
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #110.
XX
KW Antisense; ss: prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX

S. epidermidis gen
Staphylococcus aur
Staphylococcus aur
Listeria innocua C
Listeria monocyt
Bacillus lichenifo
Streptococcus poly
Genomic sequence o
Streptococcus poly
Streptococcus poly
ypHC gene of Stre
S. pneumoniae type
Streptococcus pneu
Streptococcus pneu
Streptococcus pneu
Staphylococcus aur
S. epidermidis ope
E. coli DNA for ce
DNA encoding novel
Haemophilus influe
DNA encoding Haemo
Haemophilus influe
Salmonella typhi D
Buchnera sp. genom
Haemophilus influe
S. pneumoniae derl
Genomic fragment #
C. glutamicum derl
C. glutamicum codin
C. glutamicum codin
C. glutamicum codin
S. spinosa DNA fra
S. spinosa DNA fra
Mycoplasma genital
N. meningitidis pa

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr CJ;
 XX Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 XX
 PT New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 1: Seq ID No 1463; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence is an antisense
 CC oligonucleotide of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 387 BP; 122 A; 91 C; 53 G; 121 T; 0 other:
 Query Match 100.0%; Score 387; DB 23; Length 387;
 Best Local Similarity 100.0%; Pred. No. 1.2e-98;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGTCATCTAACAGTCACCAACCTA 60
 DB 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGTCATCTAACAGTCACCAACCTA 60
 OY 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTAATAGAAATATATCA 120
 DB 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTAATAGAAATATATCA 120
 OY 121 CGTCTGTACGATTTTCATATTTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 DB 121 CGTCTGTACGATTTTCATATTTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 OY 181 ATTGTATATAAATTTGACGACGACATTTGATGCTTTGTCAATCTCTACGACGTTAA 240
 DB 181 ATTGTATATAAATTTGACGACGACATTTGATGCTTTGTCAATCTCTACGACGTTAA 240
 OY 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTAATTTTGG 300
 DB 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTAATTTTGG 300
 OY 301 TTTGGAATGTGACATCACCAATTTCAATACCAACCTGTATCAATAATTTGAATCATGTG 360
 DB 301 TTTGGAATGTGACATCACCAATTTCAATACCAACCTGTATCAATAATTTGAATCATGTG 360
 OY 361 TTAACCATTCACCTGGAAGAAATATAATAC 387
 DB 361 TTAACCATTCACCTGGAAGAAATATAATAC 387

RESULT 2
 AAF86461/c
 ID AAF86461 standard; DNA; 1311 BP.
 XX AAF86461;
 AC
 XX
 DT 26-JUN-2001 (first entry)
 XX

DE *Staphylococcus aureus* yphc coding sequence.
 XX
 XX yphc; antimicrobial; cytostatic; antiulcer; microbial infection;
 KW gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
 KW gastritis; ds.
 XX
 OS *Staphylococcus aureus*.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1311
 FT /tag= a
 FT /product= "Staphylococcus aureus yphc protein"
 XX
 XX WO200123418-A1.
 XX
 XX 05-APR-2001.
 PF 19-SEP-2000; 2000MO-US25566.
 PR 28-SEP-1999; 99US-0406968.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Zalacain M, Biswas S, Burnham MKR, Sylvester D, Mcdevitt D;
 PI Mathie TB;
 DR WPI: 2001-308138/32.
 DR P-PSDB: AAB82089.
 XX
 PT Novel yphc polypeptides of *Staphylococcus aureus* useful for diagnosing
 PT and treating microbial infections, especially infection by
 PT *Staphylococcus aureus* and *Helicobacter pylori* -
 XX
 PS Claim 2: Page 2-3; 41pp; English.
 XX
 CC The present sequence is the gene encoding yphc polypeptide of
 CC *Staphylococcus aureus*. The yphc coding sequence and protein are useful
 CC for treating and diagnosing microbial infections such as infection caused
 CC by *S. aureus* and *Helicobacter pylori*. In addition, the yphc coding
 CC sequence and protein are useful for treating diseases such as
 CC H-pylori-induced cancers, e.g. gastrointestinal carcinoma, gastric
 CC ulcers, and gastritis. The present sequence was obtained from a library
 CC of clones of chromosomal DNA of *S. aureus* in *E. coli*. The sequencing data
 CC from two or more clones comprising overlapping *S. aureus* DNAs was used to
 CC construct the present contiguous DNA sequence.
 XX
 SQ Sequence 1311 BP; 451 A; 184 C; 278 G; 398 T; 0 other:
 Query Match 100.0%; Score 387; DB 22; Length 1311;
 Best Local Similarity 100.0%; Pred. No. 1.8e-98;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGTCATCTAACAGTCACCAACCTA 60
 DB 508 GATCTTCCTCCCTCCACCAAAATGAGAAACACGTCATCTAACAGTCACCAACCTA 449
 OY 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTAATAGAAATATATCA 120
 DB 448 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTAATAGAAATATATCA 389
 OY 121 CGTCTGTACGATTTTCATATTTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 DB 388 CGTCTGTACGATTTTCATATTTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 329
 OY 181 ATTGTATATAAATTTGACGACGACATTTGATGCTTTGTCAATCTCTACGACGTTAA 240
 DB 328 ATTGTATATAAATTTGACGACGACATTTGATGCTTTGTCAATCTCTACGACGTTAA 269
 OY 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTAATTTTGG 300
 DB 268 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTAATTTTGG 209

PS Claim 27; Seq ID No 4428; 51bp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 1305 BP: 447 A; 181 C; 277 G; 400 T; 0 other;

Query Match 99.2%; Score 363.8; DB 23; Length 1305;
Best Local Similarity 99.5%; Pred. No. 1.4e-97;
Matches 385; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Y 1 GATCTTTCCTCCCTCACCACAAATGAGAAACAACGCATCTAACGAAGTCCAGCACCTA 60
DB 508 GATCTTTCTTCTCTTCCACCAAAANTGAGAAACAACGTCTAACGAAGTCCAGCACCTA 449
61 AACCATGTGACCGCTGATATCGGATACGGTTCACCAAATCCTTAATGAATGAAATCATACA 120
DB 448 AACCATGTGACCGCTGATATCGGATACGGTTCACCAAATCCTTAATGAATGAAATCATACA 389
121 CGTCTGTACGCATTTCATATTATCTACTTTGTTAACGCCATTAATGACACGGTTTTTGAG 180
Y 388 CGTCTGTACGCATTTCATATTATCTACTTTGTTAACGCCATTAATGACACGGTTTTTGAG 329
181 ATTTGTATAAAATTTGAGCGACCAATTCATGCTTGTCANACCTTCAGCAGCTTAA 240
DB 328 ATTTGTATAAAATTTGAGCGACCAATTCATGCTTGTCANACCTTCAGCAGCTTAA 269
241 CCATAAAAATATAATACATCCGCTTCATGTCATGCGCATTTCTGCTCTAATTTGTG 300
DB 268 CCATAAAAATATAATACATCCGCTTCATGTCATGCGCATTTCTGCTCTAATTTGTG 209
Y 301 TTGTGAATGTGTCATCCACCAATTTTAATACACCGCTGTATCAATAATTGAATCATGTG 360
DB 208 TTGTGAATGTGTCATCCACCAATTTTAATACACCGCTGTATCAATAATTGAATCATGTG 149
Y 361 TTAACCATTTCAACCTGAAGAATAAATAC 387
DB 148 TTAACCATTTCAACCTGAAGAATAAATAC 122

RESULT 5
AAS54865/C
ID AAS54865 standard: DNA; 1311 BP.
XX AAS54865:
DT 13-FEB-2002 (first entry)
DE Staphylococcus aureus DNA for cellular proliferation protein #1177.
XX
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
XX
PN WO200170955-A2.

[illegible]

|||||
Db 208 TTTCGAATGGTGCATCACCAATTTCATATACCACTGTATCATATATATTTGAATCATGTG 149
QY 361 TTACCATTCACCTGATACAAATTAATAC 387
Db 148 TTAACTACTACCTGATGAGATTAATTAATAC 122

RESULT 6
AAV74669/C
ID AAV74669 standard; DNA; 3621 BP.
XX
AC AAV74669;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #358.
XX
KM Computer readable medium; vaccine; S.aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FH Key location/Qualifiers
FH misc_feature 481..540
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 2281..2340
FT /tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

EP786519-A2.
30-JUL-1997.
07-JAN-1997; 97EP-0100117.
XX
PF 05-JAN-1996; 96US-0009861.
XX
PR (HUMA-) HUMAN GENOME SCI INC.
XX
PA Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 1241-1243; 3271pp; English.
XX
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock

CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
SQ Sequence 3621 BP; 1279 A; 442 C; 692 G; 1085 T; 123 other:
Query Match 84.5%; Score 327; DB 18; Length 3621;
Best Local Similarity 84.5%; Pred. No. 1.6e-81;
Matches 327; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 1 GATCTTCTCTCTCTCTACCAAAATGAGAAACAACTGATCTAACAAGTCACCAAGACCTA 60
Db 2637 GATCTTCTCTCTCTCTACCAAAATGAGAAACAACTGATCTAACAAGTCACCAAGACCTA 2578
QY 61 AACCATGTGACCCGTGATATCGATACGGTTCCACCAATCTTAATGAATGAATCATACA 120
Db 2577 AACCATGTGACCCGTGATATCGATACGGTTCCACCAATCTTAATGAATGAATCATACA 2518
QY 121 CGTCTGTACGATTTCCATATATCTACTTTGTTAAACGCTAATACGACCGGTTTTTAG 180
Db 2517 CGTCTGTACGATTTCCATATATCTACTTTGTTAAACGCTAATACGACCGGTTTTTAG 2458
QY 181 ATTTGATTAATTTGAGGACCATTTGATTCGCTTTGTCATCTTCACCGACCTTAA 240
Db 2457 ATTTGATTAATTTGAGGACCATTTGATTCGCTTTGTCATCTTCACCGACCTTAA 2398
QY 241 CCATAAAAATTAATACATCCGCTTCATCTATGCGGATTTTGCCTGCTTAATTTGCG 300
Db 2397 CCATAAAAATTAATACATCCGCTTCATCTATGCGGATTTTGCCTGCTTAATTTGCG 2338
QY 301 TTTCGAATGGTGCATCACCAATTTCATACCACTGTATCATATATTTGAATCATGTG 360
Db 2337 NNGTG 2278
QY 361 TTACCATTCACCTGATGAGATTAATAC 387
Db 2277 TTACCATTCACCTGATGAGATTAATAC 2251

RESULT 7
AAS50706
ID AAS50706 standard; DNA; 372 BP.
XX
AC AAS50706;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1930.
XX
KW Antisense; ss; prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI

XX 5. epidermidis genomic polynucleotide sequence SEQ ID NO:4072.
DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX vaccination; endocarditis; ds.
KW Staphylococcus epidermidis.
XX WO200134809-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US30782.
XX 09-NOV-1999; 99US-0164258.
XX (GLAXO) GLAXO GROUP LTD.
XX Kimmery MJ;
XX WPI; 2001-316495/33.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis -
XX Claim 8; Page 1757-1759; 2188pp; English.
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH52971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX Sequence 3269 BP; 986 A; 616 C; 451 G; 1216 T; 0 other;
SQ
Query Match 67.8%; Score 262.2; DB 22; Length 3269;
Best Local Similarity 79.8%; Pred. No. 2.1e-63;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 1 GATCTTCCTCTTCACCAAAATGAGAAACAGTCAATCAAGTCACCAAGACCTA 60
DB 2297 GATCTTCGATTCCTTTATTAAGTTTCAACAACGATCGATAGCAAAATCTCAATCTCTA 2356
QY 61 AACCATGTGACCCGATATCGGATACGTTTCAACCAATCTCAATGAATGAATATCA 120
DB 2357 ATCCATGTGAACCGAAATAGATATGATCTCCAAACCCCTAAAGAAATGAATATCA 2416
QY 121 CGTCTGATGACGATTCATATCTACTTGTAAACCGCTAATAGCGACGGTTTATG 180
DB 2417 TATCATTTAGCGATTCGAAGATATTCACCTTATTCACAGCTAATCAACAGGTTCTTAG 2476
QY 181 ATTTGATATAATTTGAGACGACATTTTCATGCTTGTGTAATCTTCACGACAGTTAA 240
DB 2477 ATTATATAAGCATTTGTCGACACATTTGCTGACCTTGTGTAAGTCTCTCAACATTTGA 2536
QY 241 CCATATAAATAATTAACATCCGCTTCATATGCGGATTTGCGCGCTCTAATTTTG 300
DB 2537 CCATATAAATGATGACATCTGCTTCTCAATTTGCTGCTGACGAGATTTGAG 2596

QY 301 TTGGATGTCGATCACCATTTCATATACCACTGTATCATATATTTGAAATCATG 360
DB 2597 TTGAAAAGGAGCATCTTCCATTTCAATACCACTGTATCATATGTTAAATCATG 2656
QY 361 TTACCATTTACCTGAGAAATTAATAC 387
DB 2657 TTACCACTCGCGAGATGAATTAATAC 2683
RESULT 10
AAS50205
ID AAS50205 standard; DNA; 298 BP.
XX AAS50205;
AC AAS50205;
XX 13-FEB-2002 (first entry)
DT Staphylococcus aureus cellular proliferation inhibitory sequence #1429.
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1429.
XX Antisense: ss; prokaryotic cellular proliferation;
XX antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX Claim 1; Seq ID No 2782; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms.
XX oligonucleotide of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp://ipo.int/pub/published_pct_sequences.
XX Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;
SQ

Query Match 60.5%; Score 234; DB 23; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e-56;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 60
|||||
65 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 124
|||||
61 AACCATGTGACCCCTGATTCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 120
|||||
125 AACCATGTGACCCCTGATTCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 184
|||||
121 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACCGCTAATACGACCGTTTGTAG 180
|||||
185 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACCGCTAATACGACCGTTTGTAG 244
|||||
181 ATTTGTATATAATTTGAGGACCATTTTCATCGTTTGTGTCATCCTTCACGCA 234
|||||
245 ATTTGTATATAATTTGAGGACCATTTTCATCGTTTGTGTCATCCTTCACGCA 298
|||||

RESULT 11
AAS50723
ID AAS50723 standard; DNA; 298 BP.
AC AAS50723;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1947.
XX
XX Antisense; ss: prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207127P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR
XX
XX
PS Claim 1: Seq ID No 3300; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery

CC Programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;

Query Match 60.5%; Score 234; DB 23; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e-56;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 60
|||||
65 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 124
|||||
61 AACCATGTGACCCCTGATTCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 120
|||||
125 AACCATGTGACCCCTGATTCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 184
|||||
121 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACCGCTAATACGACCGTTTGTAG 180
|||||
185 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACCGCTAATACGACCGTTTGTAG 244
|||||
181 ATTTGTATATAATTTGAGGACCATTTTCATCGTTTGTGTCATCCTTCACGCA 234
|||||
245 ATTTGTATATAATTTGAGGACCATTTTCATCGTTTGTGTCATCCTTCACGCA 298
|||||

RESULT 12
AB067194/C
ID AB067194 standard; DNA; 319630 BP.
XX
XX AB067194;
XX
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua config DNA sequence #7.
XX
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
XX Listeria innocua.
OS
XX
PN WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
DR
XX
XX
PS Claim 5: SEQ ID 7; 180pp; French.

The present invention relates to nucleic acid sequences (AB067188-AB071212) from Listeria sp. The sequences are useful as probes and primers for identification and/or detection of Listeria (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be

CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of *Listeria* (potential therapeutic agents), also for
CC treating infections by *Listeria*, and are useful as immunogens in
CC anti-*Listeria* vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at [ftp://wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 6 other:

Query Match 53.2%; Score 205.8; DB 24; Length 319630;

Best Local Similarity 70.9%; Pred. No. 5.5e-47;

Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 3 TCTTCTCCTCTTCCACCAAAATGAGAACAACTGCATCTACAGTCACCAAGCCTTAA 62

DB 51123 TCCTTCTCTCTTTGGAAAATGACACGACAGATCAAGATTCACCAAGCCTTACT 51064

QY 63 CCATGTGACCCCTGATATCGGATACGGTTCCACCAAACTCTAATGAATAGAAATCAATACAG 122

DB 51063 CCATGAGAACCCAGAAATGGATACGGCTCACCAACCCAGAAATGAATCAATATATAC 51004

QY 123 TCTGTACCATTTCCATATATATCTACTTTGTTAACCGCTATATACAGCGGTTTTAGAT 182

DB 51003 TGAATCCGATTTCTGGGTTATCTACTTTATTAATGCTAAACAAATGGTTTATAGAC 50944

QY 183 TTGTATAAATTTGAGCGACATTCATCGCTTGTGTCATCGCTTACAGCAGTTAAC 242

DB 50943 CGGTAAAGAAATTTTGTCTACTTGTCTGTCATGGTACCCCTTACAGCAGTTAGTA 50884

QY 243 ATAAATATATATACATCGCTTCATCTATGCGGATTTTCCTCGGCTCTAATTTGTGT 302

DB 50883 ATAAATATATATACATCGCTTCATCTATGCGGATTTTCCTCGGCTCTAATTTGTGT 50824

QY 303 TGGATGTGTGATCACAATTTTCATATACCGCTGATCAATATATGAAATCATGTGT 362

DB 50823 AAGATGTGTGCTCGGAAGATCAATACCACTGATCAATATGTTAATTTCTTTCCA 50764

QY 363 AACCATTCACCTGAGAAATATATAC 387

DB 50763 ACCCATTCGCTGAAATATATATGTC 50739

RESULT 13

ABA03041 ID ABA03041 standard; DNA: 2944528 BP.

XX ABA03041;

DT 05-FEB-2002 (first entry)

XX *Listeria* monocytogenes EGD-e genome sequence.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KM vitamin B12; bacterial infection; disease; ds.

XX *Listeria* monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Eshti H, Dehoux P;

PI Dusserret O, Chetoui F, Nedjati H, Glaeser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for *Listeria* monocytogenes, useful e.g. for treatment

XX and prevention of *Listeria* and related bacterial infections, and

XX related polypeptides

PS Claim 1; SEQ ID No 1; 192pp; French.

XX The present sequence is the genome sequence of *Listeria* monocytogenes

XX EGD-e. This sequence and fragments of this sequence are useful for

XX selecting probes and primers for detecting genes in *L. monocytogenes* and

XX related organisms, and to study genetic polymorphisms and other genomes.

XX Proteins (AB847297-AB850149) expressed from the present sequence are

XX useful for raising specific antibodies, identification of *L.*

XX *monocytogenes* and related organisms, and for biosynthesis and

XX biodegradation, especially biosynthesis of Vitamin B12. This sequence and

XX proteins encoded by it are also useful for selecting compounds that

XX regulate gene expression and cell replication and modulate *L.*

XX *monocytogenes*-related diseases. In addition, this sequence and proteins

XX encoded by it are useful in pharmaceutical and vaccines compositions for

XX the treatment or prevention of infections by *L. monocytogenes* and related

XX organisms.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at [ftp://wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other:

Query Match 53.2%; Score 205.8; DB 24; Length 2944528;

Best Local Similarity 70.9%; Pred. No. 1.1e-46;

Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 3 TCTTCTCCTCTTCCACCAAAATGAGAACAACTGCATCTACAGTCACCAAGCCTTAA 62

DB 2011933 TCCTTCTCTCTTTGGAAAATGAGCGGACAGATCAAGATGCGCAAGCCTTACT 2011992

QY 63 CCATGTGACCCCTGATATCGGATACGGTTCCACCAAACTCTAATGAATAGAAATCAATACAG 122

DB 2011993 CCATGGAACCCAGAAATTTGATACGGCTCACCAACCCAGAAATGAATCAATATAT 2012052

QY 123 TCTGTACCATTTCCATATATATCTACTTTGTTAACCGCTATATACAGCGGTTTTAGAT 182

DB 2012053 TGAATCCGATTTCTGGGTTATCTACTTTATTAATGCTAAACAAATGGTTTATAGAC 2012112

QY 183 TTGTATAAATTTGAGCGACATTCATCGCTTGTGTCATCGCTTACAGCAGGTTAAC 242

DB 2012113 CGGTAAAGAAATTTTGTCTACTTGTCTGTCATGGTACCGTAACTTCTTACAGCAGGTTAGTA 2012172

QY 243 ATAAATATATATACATCGCTTCATCTATGCGGATTTTCCTCGGCTCTAATTTGTGT 302

DB 2012173 ATAAATATATATACATCGCTTCATCTATGCGGATTTTCCTCGGCTCTAATTTGTGT 2012232

QY 303 TGGATGTGTGATCACAATTTTCATATACCGCTGATCAATATATGAAATCATGTGT 362

DB 2012233 AAAAATGTTCATCGGAAGATCAATACCACTGATCAATATATGTTAAATCTTTTCCA 2012292

QY 363 AACCATTCACCTGAGAAATATATAC 387

DB 2012293 ACCCATTCGCTGAAATATATATGTC 2012317

RESULT 14

AB069245 ID AB069245 standard; DNA: 3011208 BP.

XX AB069245;

XX 29-AUG-2002 (first entry)

```

DE Listeria innocua DNA sequence #684 .
XX
XX Antibacterial: Listeria: food contamination; mutational analysis:
XX infection; ds.
XX
XX Listeria innocua.
OS
XX MO200228891-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR03061.
XX
XX 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
XX
XX WPI: 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators -
XX
XX Claim 5; SEQ ID 2058; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (AB067188-AB071122) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO atftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 3011208 BP: 941651 A; 568176 C; 559189 G; 942192 T; 0 other;
SQ
Query Match 53.2%; Score 205.8; DB 24; Length 3011208;
Best Local Similarity 70.9%; Pred. No. 1.1e-46;
Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 3 TCTTCTTCTCTTCCATCCAAATGAGAAACACATGATCTAACAGTCCACCAAGCTTAA 62
DB 2066305 TCCTCTTCTCTTCTTTGGAAAAATGAGCAAGAACAGCATATACATACCAAGCCCTTAGT 2066364
QY 63 CCATGTGACCTGATATCGATAGCGTTCCACCAATCCTAATGAAATGAAATCATACAG 122
DB 2066365 CCATGAGAACAGAAATTTGATAGCGGTCCACCAACCAAGAGAAATAAAGTCATTAATC 2066424
QY 123 TCTGTAGCCATTTTCCATATTAATCTACTTTGTTAAACCGCTAATACGACCGGTTTTAGAT 182
DB 2066425 TGATCCACCATTTCTGGGGTATCTACTTATTAATCCGTAACAAATGCGTTTATTGAC 2066484
QY 183 TTGTATAAATTTGAGCGACCATTTTCATCGCTTTGTGTGCATTCCTTACGACGCTTAAC 242
DB 2066485 CGGTAAACAAATTTTGTACTTGTGTCGTCGATCGGTAAACCCCTTACGACCAATGGA 2066544
QY 243 ATAAAAAATAATACATCCGCTTCATATCTATGCGAATTTTGCCTGCGCTAATTTGAGT 302
DB 2066545 ATAAAAAATAATACATCTGCTTCATCAATTTGCGATTTCCGTTGTGCGCGAATTTGCT 2066604
QY 303 TGAATGTGATCACCACAAATTTCAATACACCTGTATCAATATAATGAAATCATGTGTT 362
DB 2066605 AAGAAATGTGCGGAGAAAGATCATACACCTGTATCAATATAATGTTAAATTTCTTTCCA 2066664
QY 363 AACCATTTCACTGAAGATTAATAC 387
1 | 111111 | 11111 | 111 | 1 |

```

DB 206665 AGCCATTCGGCTGATTAATATATGC 2066689

RESULT 15
ID ABR75008/C
XX ABR75008 standard; DNA; 1311 BP.
AC ABR75008;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #2299.
XX
KM Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
XX physiological provocation; ds.
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
XX 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
XX (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
PI WPI; 2002-416684/44.
DR
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
PS Claim 4; SEQ ID NO 2299; 200pp; English.

XX
XX The invention describes a method of monitoring differential expression of
XX genes in a first Bacillus cell relative to expression of the genes in
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX isolated from Bacillus cells to a substrate containing array of Bacillus
XX genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
XX a spot in the array. The method is useful for measuring the expression of
XX genes in a first Bacillus cell relative to expression of the same genes
XX in one or more second Bacillus cells. The method is useful for monitoring
XX global expression of several genes from a Bacillus cell, discovering new
XX genes, identifying possible functions of unknown open reading frames and
XX monitoring gene copy number variation and stability. Monitoring changes
XX in expression of genes may be used to provide a representation of the way
XX in which Bacillus cells adapt to changes in culture conditions,
XX environmental stress or other physiological provocation. Extensive
XX follow-up characterisation is unnecessary, when one spot on an array
XX equals one gene or one open reading frame, since sequence information is
XX available. This sequence represents a genomic sequence tag (GST) used in
XX the method of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1311 BP; 376 A; 279 C; 353 G; 303 T; 0 other:

Query Match 45.1%; Score 174.4; DB 24; Length 1311;
Best Local Similarity 67.1%; Pred. No. 6 le-39;
Matches 247; Conservative 0; Mismatches 121; Indels 0; Gaps 0

00 AAAATGAGAAACACTGCACTTCAACAGTCAACCAAGACCTTAACCATGTGACCTTGATAT 79

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 ; Search time 1580.02 seconds
(without alignments)
10020.134 Million cell updates/sec

Title: US-09-815-242-1463

Perfect score: 387

Sequence: 1 gatctctctctcttcacca.....ttcactgaagaataaac 387

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_in:*

32: em_htg_inv:*

33: em_htg_other:*

34: em_htg_mus:*

35: em_htg_pla:*

36: em_htg_rnd:*

37: em_htg_man:*

38: em_htg_vtl:*

39: em_sy:*

40: em_htgo_hum:*

41: em_htgo_mus:*

42: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	387	100.0	1308	6	AX622668
2	387	100.0	301550	1	AP003134
3	387	100.0	333750	1	AP004827
4	387	100.0	346900	1	AP003362
5	262.2	67.8	3269	1	AF270032
6	262.2	67.8	3269	6	AX145350
7	262.2	67.8	300892	1	AE016747
8	210.6	54.4	290117	1	AE017028
9	209	54.0	304680	1	AE017002
10	205.8	53.2	313450	1	AL596170
11	205.8	53.2	319630	6	AX413016
12	205.8	53.2	347050	1	AL591981
13	205.8	53.2	349980	6	AX417046
14	205.8	53.2	349980	6	AX641670
15	198.4	51.3	24887	1	BACSERA
16	198.4	51.3	213680	1	BSUB0012
17	191.4	49.5	302173	1	AE016951
18	188.2	48.6	300050	1	AP004599
19	185	47.1	300550	1	AP001512
20	182.4	47.1	302050	1	AL935257
21	174.4	45.1	1311	6	AX433884
22	162.6	42.0	11071	1	AE006309
23	162.6	42.0	12434	1	AE006498
24	162	41.9	1311	6	AX607165
25	162	41.9	44145	6	AX602195
26	162	41.9	174050	1	SAG766852
27	160.4	41.1	20601	1	AE014265
28	159.4	41.2	52276	1	AE014141
29	159.4	41.2	323825	1	AP005146
30	157.8	40.8	12370	1	AE009978
31	153	39.5	12540	1	AE008523
32	151.4	39.1	1308	6	AX570314
33	151.4	39.1	5066	6	AX218847
34	151.4	39.1	5066	6	BD003759
35	151.4	39.1	10310	1	AE007464
36	151.4	39.1	151947	2	SPNEU1902
37	151.4	39.1	349980	6	AX571764
38	151.4	39.1	349980	6	AX571765
39	149.8	38.7	13860	1	AE015016
40	148.2	38.3	3737	1	AB016077
41	139.4	36.0	301278	1	AE015939
42	125.2	32.4	296750	1	AP003191
43	117	30.2	10861	1	AE007680
44	109.2	28.2	3557	1	AY094626
45	101	26.1	960	6	AX144037

ALIGNMENTS

RESULT 1
AX622668/c
LOCUS AX622668 1308 bp DNA
DEFINITION Sequence 5631 from Patent WO02094868.
ACCESSION AX622668
VERSION AX622668.1 GI:28450653
KEYWORDS
SOURCE
ORGANISM
Staphylococcus aureus
Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
AUTHORS
TITLE Masignani,V.C., Mora,M.C. and Scarselli,M.C.
JOURNAL
Patent: WO 02094868-A 5631 28-NOV-2002;
Chiron Spa (IT)

FEATURES
source
Location/Qualifiers
1..1308
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"

BASE COUNT 450 a 184 c 278 g 396 t
ORIGIN

Query Match 100.0%; Score 387; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCTCTTCCTCCCTTCACCAAAATGAGAAACACATGCATCTACAGTCACCAAGACCTA 60
|||||
Db 508 GATCTCTTCCTCCCTTCACCAAAATGAGAAACACATGCATCTACAGTCACCAAGACCTA 449
|||||

QY 61 AACCATGTGACCCGTATATCGGATACGGTTACCAAAATCCATATGAATGAATCATACA 120
|||||
Db 448 AACCATGTGACCCGTATATCGGATACGGTTACCAAAATCCATATGAATGAATCATACA 389
|||||

QY 121 CGTGTGATGACATTTTCATATTTACTTTGTAAACGCTAATACGACGGTTTTTTAG 180
|||||
Db 388 CGTGTGATGACATTTTCATATTTACTTTGTAAACGCTAATACGACGGTTTTTTAG 329
|||||

QY 181 ATTTGTATTAATAATTGAGCGACCATTTTCATCGCTTTGTGTCAATCCTTCACGACGTTAA 240
|||||
Db 328 ATTTGTATTAATAATTGAGCGACCATTTTCATCGCTTTGTGTCAATCCTTCACGACGTTAA 269
|||||

QY 241 CCATATAAATAATATACATCGCTTCATGCGATTTTCGCTGCGCTTAATTTTG 300
|||||
Db 268 CCATATAAATAATATACATCGCTTCATGCGATTTTCGCTGCGCTTAATTTTG 209
|||||

QY 301 TTGGAATGTGTCATCACCATTTTCATATACACCGCTATCATTAATATTGAATCATGTG 360
|||||
Db 208 TTGGAATGTGTCATCACCATTTTCATATACACCGCTATCATTAATATTGAATCATGTG 149
|||||

QY 361 TTAAACCATTCACCTGAAGAATAATAC 387
|||||

Db 148 TTAAACCATTCACCTGAAGAATAATAC 122
|||||

RESULT 2
AP003134 301550 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE Bacteria; Firmicutes; Bacillales; Staphylococcus.
AUTHORS
1
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Baba,T., Yuzawa,H., Kobayashi,I.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
TITLE Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center, 2chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,

COMMENT
FEATURES
source
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
On Jun 12, 2001 this sequence version replaced gi:13701258.
Location/Qualifiers
1..301550
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="genomic DNA"
/strain="N315"
/sub_species="aureus N315"
/db_xref="taxon:15879"
complement(249..947)
/gene="SA1293"
complement(249..947)
/gene="SA1293"
/note="ORFID:SA1293"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42553.1"
/db_xref="GI:13701259"

/translation="MSIFMTIIVVILNVLPLAOKHKVSNYKYSQVRSKTKGR
EVALDILHANGIYDEVEVKGEGFLTDHPKKRVLSLPANDRPSVAGTAAAEVG
HAIDHOGYWFLEFRALVLPVNLGSSYSYMIIMGILITAIAGSFGSTALMIGM
SLAVFESIYTLPEFEDASSRAWKQITALINVEREKYKARKVLSAAATVYAAVAV
AELVRIILARSDD"
complement(1001..1588)
/gene="SA1294"
complement(1001..1588)
/note="ORFID:SA1294"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42554.1"
/db_xref="GI:13701260"

/translation="MTINFMQYTLTQORSWLMMLICNLGMIVGYWGEQLSHTPM
HKRIPDPSPTALIVLISILTIQKONSIIIEALAFVTLFYGIAVYIMNLFIIEO
GDITNGIYIMSHSISMAVOAYFPYPRKRSWIGTSVAMIWFLNDYIDFHLQPPY
DFTTHVMOIGVLSCLSYFGLIYIELNKLKCK"
complement(1578..2153)
/gene="SA1295"
complement(1578..2153)
/note="ORFID:SA1295"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42555.1"
/db_xref="GI:13701261"

/translation="MSETLNQIKESFIEYLFOYRKSRIAVWLVNLYIKVNAKLANI
HFVDRKINHHEPLETAEGVSHASAIQFTKRNIKLNNTNEIDYIANNHCARDIQTHA
NVSKREQRLLDLIVAQLESPSYQYTLADLNMAIDRRKHALLDYLLHNIDLSIQM
EKORFYOLTOILNTLKLNVKHNQFEDLADD"
complement(2167..3411)
/gene="SA1296"
complement(2167..3411)
/note="ORFID:SA1296"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42556.1"
/db_xref="GI:13701262"

/translation="MEDIYKLLIDINLOKLENLDSRVBEAITTDNDALFILGETLYN
GLMPQGLEVERFVLYHKYDESELIVFIEGLMSNQDEALEVLSYVPSPEKLMLE
ADLYQOINMMEYALIDQLEALEPNDEPIIHFALEMYGDOYLRASSEVETLETG
EYQVNGVLFSSMADCSLOSNGVSAIRLYDEINDENMTSEDYLRKASYSOKNDTQE
AIKINFTLLSKDQDYIOGYLYLYSLENKYNPDAIEGKGLRISQPKELMYTTCG
LEIEGDNBEYLLKQALEVDNAYOEPLILSDYIRNEDEYALIEITLTYVDEEDLD
PTEFMWHLAFAYGQEEERKEQHFELAYPTMKTINIDMSDYIFYLIEIGKEQATITL
NQLLEPSSNEMWHDLSRLQS"
complement(3418..4716)
/gene="aroc"
complement(3418..4716)

```
/gene="aroA"
/feature="ORFID:SA1297"
/codon_start=1
/transl_table=11
/product="3-phosphohikimate 1-carboxyvinyltransferase"
/protein_id="Bab42557.1"
/db_xref="GI:13701263"
/translation="MVEQIIDISGPLKEIYEPGDKSMTHRAIMLASLAEGTSNIX
PLIGEDCRRTMDIFRLGVDIKEDDEKLVNNSPGYKAFKPHQVLYGNSGTTLLA
GLISGIESISGDVSGIKRPMRVRLPKLMDANEGIEDNMTPLITKPSYIKGN
XOMEVSAQYKSAILFRASLESNTTYIKELDVSNNHTETFRHNPIDEKRSITTT
PDAIHIKPADFHPGDISAFAFIVAALITPESDVTIHNVGINPRSGIIDIVKRG
GNIQLNQTGAPPTASIRIQYPMLOPIITIEBELPKAIDELPVIALICTQAVGST
IKDAEIKYETNRIDTADMLNLGFELOPTNDGLIHPSEKTAATVDSLDRIG
MMLAVASLSSEPVKIKOPDVAIVVSFPGFLPKLLENEG"
complement(4726..5790)
/gene="arob"
/complement(4726..5790)
/gene="arob"
/feature="ORFID:SA1298"
/codon_start=1
/transl_table=11
/product="3-dehydroquinate synthase"
/protein_id="Bab42558.1"
/db_xref="GI:13701264"
/translation="MKIQTYPSNNYPIFEVGHAIHISTYIDQFDSFLIDENHNO
YFADKFDLISYENHVKVILPAGEKTFEYOGETLEYILSHVYTRMTATIAVGGAT
GDFAGFYATLLRGVHPHIOVPTIILHSDSVGKVGKSGKGLICAFYRPAVYTD
LDFLKLTPFQILSGYAEVYKHALNLESTTOETEOGFKREILQSGMGRKTIKAI
ETKLDIVADEKQGVKFLNLGHTFGHAEVYHKKIAHGAVMIGIILYQIVANILEN
SNHDIQHYVLYKRGVPLETYIDIDFETIYOYMLSDKNDKQGVOMVLIKHGDIYV
QHIDQTLQHACQDLKTYEK"
complement(5816..6982)
/gene="aroc"
/complement(5816..6982)
/gene="aroc"
/feature="ORFID:SA1299"
/codon_start=1
/transl_table=11
/product="chorismate synthase"
/protein_id="Bab42559.1"
/db_xref="GI:13701265"
/translation="MRYLTSGESHGPOLVIVGIPANLEIKVEDINKENFKRGYGG
KGRMOIEXDTEIVISGVNRYGLSPITVWVNDPTHRKINGADPISEEBEENK
RTTKPRPGHADVGVKMYNHRDLNRYLESSARETARAVAGLCVILQOIDDY
SRVEIGIGIKDQFVDSSEFKANLDRNDVVIDSIAQMRDKIDEKNGSDIGYV
QYVENNRPVGVSYVHDKLKDIAQGVISNAFKVSGEGFKAEKSGEIQDEI
LVSEIGYRGSNHLGLEGSGMSGMPTIVNGVMPITPLKPLNSVDINTKEDFKAT
IERSDCAVPAASIVGEHVVAFEIARALBEPQSNHIELOQOITADRRQLNVEF"
complement(7368..7568)
/gene="SA1300"
/complement(7368..7568)
/gene="SA1300"
/feature="ORFID:SA1300"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="Bab42560.1"
/db_xref="GI:13701266"
/translation="MKNISIIIGIIMFLKLMELLYGAIFLDRPLNDITKIIFLTLI
YIPYLVAKELIIFLKSRYNSA"
complement(7597..7722)
/gene="SAS045"
/complement(7597..7722)
/gene="SAS045"
/feature="ORFID:SA5045"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="Bab42561.1"
/db_xref="GI:13701267"
/translation="MWLVIMIHKNARFLCCSKLLEKYSFLVTRIKKRGIFIK"
complement(7777..8226)
```

```
CDS
/gene="ndk"
/complement(7777..8226)
/gene="ndk"
/feature="ORFID:SA1301"
/codon_start=1
/transl_table=11
/product="nucleoside diphosphate kinase"
/protein_id="Bab42562.1"
/db_xref="GI:13701268"
/translation="MERTFLMKIPDANORNLIGEVISIRERKGLKLVGKLMQVPMEL
AETHYGEQKRPYNDLISFTSAHYFAMVVEGDVAIVSRHIIIGSTNPESPSGIR
GDGLITVGRNIHGSDSLESAREINLMFNENETSYASPRDAWLYE"
complement(8318..9277)
/gene="gerCC"
/complement(8318..9277)
/gene="gerCC"
/feature="ORFID:SA1302"
/codon_start=1
/transl_table=11
/product="heptaprenyl diphosphate synthase component II"
/protein_id="Bab42563.1"
/db_xref="GI:13701269"
/translation="MAKLMMNNEIKKVEQRLEKAIKSKDSVLEQASLHLSSGGKRV
PAFVILSSQFGKDEQTSQRTYQVAVALELLHMAITLVHDVIDSDKRRKLTLSKMD
```

Query Match 100.0%; Score 387; DB 1; Length 301550;
Best Local Similarity 100.0%; Pred. No. 1 2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GATCTTCTCTCTTCACCAAAATAGAAACACGATCTAACAGTCACCAAGACCTA	60
Db	13271	GATCTTCTCTCTTCACCAAAATAGAAACACGATCTAACAGTCACCAAGACCTA	13330
QY	61	AACCATGTGACCCCGATRTGGATAGCGTTCCACCAATCCATAGATAGAAATCATACA	120
Db	13331	AACCATGTGACCCCGATRTGGATAGCGTTCCACCAATCCATAGATAGAAATCATACA	13390
QY	121	CGTCTGACGATTTCCATATATCTACTTTGTTAAACCGCTAATACGACGGTTTTTAA	180
Db	13391	CGTCTGACGATTTCCATATATCTACTTTGTTAAACCGCTAATACGACGGTTTTTAA	13450
QY	181	ATTTGTATAAATTTGAGCGACATTTTCATCGCTTGTGTCATCTTCACGACGTTAA	240
Db	13451	ATTTGTATAAATTTGAGCGACATTTTCATCGCTTGTGTCATCTTCACGACGTTAA	13510
QY	241	CCATTAATAATTAATACATCCGCTCATCTATAGGCGATTTTCGCTGCGCTTAATTTGTG	300
Db	13511	CCATTAATAATTAATACATCCGCTCATCTATAGGCGATTTTCGCTGCGCTTAATTTGTG	13570
QY	301	TTTGAATVGGTGCATCAACATTTCAATACACCGCTGTATCAATATATTTGAATCATGTG	360
Db	13571	TTTGAATVGGTGCATCAACATTTCAATACACCGCTGTATCAATATATTTGAATCATGTG	13630
QY	361	TTAACCATTCACCTGAAGAATAAATAC	387
Db	13631	TTAACCATTCACCTGAAGAATAAATAC	13657

RESULT 3
AP004827 333750 bp DNA 1linear BCT 24-Apr-2003
LOCUS
Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 6/10.
ACCESSION
AP004827 BA000033
VERSION
AP004827.1 GI:21204509
SOURCE
Staphylococcus aureus subsp. aureus MW2
ORGANISM
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1 Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
Negai,Y., Iwama,N., Asano,K., Naito,T., Kuroda,H., Cui,L.,
Yamamoto,K. and Hiramatsu,K.

TITLE	Genome and virulence determinants of high virulence community-acquired MRSA
JOURNAL	Lancet 359 (9320), 1819-1827 (2002)
MEDLINE	22040717
PUBMED	12044378
REFERENCE	2 (bases 1 to 333750)
AUTHORS	Director-General, Biotechnology Center, Aoki,K., Oguchi,A., Nagai,T., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2chome 48-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bioente.go.jp, URL:http://www.bio.nite.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES	Location/Qualifiers
source	1. .333750
	/organism="Staphylococcus aureus subsp. aureus MW2"
	/mol_type="genomic DNA"
	/strain="MW2"
gene	/db_xref="taxon:196620"
	complement(278. .619)
gene	/gene="MW1341"
	complement(278. .619)
CDS	/gene="MW1341"
	/note="ORFID:MW1341"
	/codon_start=1
	/transl_table=11
	/product="hypoetical protein"
	/protein_id="BAB95206.1"
	/db_xref="GI:21204510"
	/translation="MTNITKEVFDNLDEIDVFAKNKTLGSEAKPYLDEYHSKIIDYFKVNDIGNIDFDLNOYPVAVNFKERYDMIERKHYFMGYRQMKTEKTELKMA SYQRLKNKQV"
gene	complement(624. .1283)
	/gene="nth"
CDS	complement(624. .1283)
	/gene="nth"
	/note="ORFID:MW1342"
	/codon_start=1
	/transl_table=11
	/product="endonuclease-like protein"
	/protein_id="BAB95207.1"
	/db_xref="GI:21204511"
	/translation="MWSKKALMIDVIANNPPDAECLEKHDNPELTIAVLNACQTDVLVAVRTTELFEKTKTPEDYLAVSDELMNDIRSIGYRNKAKNIKRLCOSLIDQFN GEIDPTNKELESLAGVGRKTANVMSVAFDEPSLAVDTHVERVSRLCINMKDVAHQ VEDRLCSVIPRDMNRSHQDLFFGRYHCLARKPKCDICPLLEDCREGQKRYKASLKE A"
gene	complement(1273. .1959)
	/gene="MW1343"
CDS	complement(1273. .1959)
	/gene="MW1343"
	/note="ORFID:MW1343
	hypoetical protein, similar to chromosome replication initiation protein dnad"
	/codon_start=1
	/transl_table=11
	/protein_id="BAB95208.1"
	/db_xref="GI:21204512"
	/translation="MDKYQLKAPVIVIRELLDHSYDGLDEODLVILLKLIVASETS NKPSIELQKSGMQPRDITWVIONLQRELELQYOKDEGRTEYNDPPEKLSHILKQSMETKEONSKERKQLEFVLEDTFARPLSPYEITLQWIDVDHDSATIQ AALDEANLNLSEFYMDRILLNMKKNVKTIDSRKIREKFNKPMHTVTKYKFD WLNGLNLDGK"
gene	complement(2287. .3579)
	/gene="asnS"
CDS	complement(2287. .3579)
	/gene="asnS"
	/note="ORFID:MW1344"
	/codon_start=1
	/transl_table=11
	/product="asparaginyl-tRNA synthetase"
	/protein_id="BAB95209.1"
	/db_xref="GI:21204513"
	/translation="MKTIKQAKDHLNDVYTGAWLTNRSSGKIAFLQLRDGTGFMQGVVXSEVDEEVEFLAKEITQESSLYVTGITEEDNRSDLGEMQKSEIVSEADYPTPKNGHTEFLMDHRLMLRSKQKHAVMKIREVYRATYEFNFKVDPIDPLTAS APEGTSLEPHKYFPQDAPISQSGOLYLEAAMAMGAKVPSGPGFRASKSTRHLIE FMWITGEMAFTHAASLEIOEDYVTHVYKSVLENKLEKITERSTSLKELYANFPFR ISYDAIEFLKSGFDDLEWGEDFAPHELTAIANNYDLVPTFTNTPTKIKPYQPNP ENEEFTVLCADLIAPEGGEIIGGSEKVDLELLEQVRKEHGLEDEEAYSYVDLRRYGS VPHCGFLGLEETVYAMISGVEHVRETAFFPRLNRLP"
gene	complement(3901. .6594)
	/gene="ding"
CDS	complement(3901. .6594)
	/gene="ding"
	/note="ORFID:MW1345"
	/codon_start=1
	/transl_table=11
	/product="probable Arp-dependent DNA helicase ding"
	/protein_id="BAB95210.1"
	/db_xref="GI:21204514"
	/translation="MGMAVYAVDLETTGNQDLPDDIIOIGITPVFNNOIIDTVHSMIRTNLEIPEIQTALSTIEENMLQOAPYFNQVQOEDYDKIKDCIFVANHVDPLNFKKAFKDCNIOYRPPKKVIDTLEIFKIAFPDTSYQSELAENHGTALANHRADEDAATTAKLMLIAPEKPEKLPDITLKOIYYLSKQLYDLDYDIFEEVROYDAKPLDKSEKPEOIT YRKQYDFKPTTNVNGSLKSLSKAVDQGLTYRQOYLAEPTIIDOLMHSKAMIEA SLGSKSLAYLALALMNIETGKHVISTNRKLQSLDEKIDIPMANALNFKIKAL ILKSDYISLGLISQILKDDTSNVEVNLKQQLMTETPEBGDIQELNLGGQKMFY DQKIETVYPARHDVYVNFIRKNAONIOIGITHNALHSDVENSIYOLFDCIYDEA HRLPYALNOVYTNELSYADIKQGLIGKNEKEKLKALIDQLEKQRIIEKLDIADIDI FGKLSMNEIHELNEOLFSTFTTIIINDSGVDDDIHRPHNVFTFETKQILMDLHAIID KLNRLEIFNGLSHKTVKSLKQLLYLIDKRNKEIOSKAGHTSPISIKNLSSKSTIR LTVQYAAKDVLTQVLEKFSNLITSGTLRNFSEFAKQLEKNDVHFNTEVNTSL QSAKNTSVFIPSDVASYOYKNIDEVASIVSIIEYTTISKCLVFTSKMMHMQ DMLNELPEFEDVYVLTQONQNYKIVQGFNNEDKALIGTSTFGFPGFQANGIKCVI IAKLPEFNKHNAAKYWLMDSEFTSEKFEVLLPDVAVRPFQGRILIRSENDGIIYSFD DRLINSYKKNFEQRTLENYROKKQIDQFGKLLRQIQKKKK"
gene	complement(6618. .7589)
	/gene="MW1346"
CDS	complement(6618. .7589)
	/gene="MW1346"
	/note="ORFID:MW1346"
	/codon_start=1
	/transl_table=11
	/product="truncated hypoetical protein"
	/protein_id="BAB95211.1"
	/db_xref="GI:21204515"
	/translation="MSKYSDVYLQLLYKKNPNYISGOSIASLNSIRYAVKVIDQLK LEGCRIDSVNKHGHLIQQLPDIWYOGIIDQYTKSSALPDSFSEVYDSIDSTOLAARSL VGNOSSEFLISDEQTKGRGFRMRHSSSGGGLMMSVYLRPVASMSIKNFLALG IRDAIQHFSQDEPKYKMPNDIYIDNGKYGCGLTENVVANNQDIEALIGGIGINTQOLE NFDESIRHRATSVQLDKNKLDRYOFRLRLQLEIKRKNQFLTPESEIREBYIAASN IWNRTLLETTENDKQFGQALDLDYDGYLIVDEAGESHRLISADIDF"
gene	complement(7576. .8778)
	/gene="MW1347"
CDS	complement(7576. .8778)
	/gene="MW1347"
	/note="ORFID:MW1347
	hypoetical protein, similar to poly(A) polymerase"
	/codon_start=1
	/transl_table=11
	/protein_id="BAB95212.1"
	/db_xref="GI:21204516"
	/translation="MDKSLPEQARPILQIODNGFEAYVVGSAVDYVYMGRIHIDIT ITTSAPDEIESIFSTIIPVGRKEHTINVPDENDEVTTFRAEDYVYHRRPSGVTV RDLVEDLQDRDFTNMAIAMDAKYDYDFQCOIDNNIRITRYGAEERFQEDALMRI RCLRFQSOISFDIATEPEARMIOAWDKIPISIEIYELTKLMGINKVESFNPLKS LKAENYMPYFELIDMNOINVTAEIDLELLIAYSVKPDINYSIKPLKISNOVDIND YIOIKNALPSITTKEQLKMFYVDYDTHLIKVNVAADYIKANDIOGHEPLVYNQOTID ETLHRLPMMHNRDMVNGCVLMAHLNANSGPWLADVLQIETAIYTGAVSNEETELK WVDNHVKI"
gene	complement(8783. .9925)
	/gene="MW1348"
CDS	complement(8783. .9925)

```
/gene="WM1348"
/Note="ORFID:WM1348
hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"
/translation="MKIGTCYPSMGSGGIATTELGIKLAERGHVHTTSNIPRIR
KPLPMIFHOVEVNOYAVFOYPPYDITSTKTAEVKKEGDLIMHYAVPAICIL
AREMSGRDKIKMTTHAGTDITVYLGDHSLQGIKIKESDVIYTSVSLAODETHI
ETNKEIPIYVRENEPPTKHTALKSQFGLADEKVLIHVSNFQVKRIDTIIETE
AKVREKIPSKLLDGDPELVPMQRLTEVNEQVLELGKQDCSEFYQLDVLLE
SEKESFGJLTLEAMKTGVPIGSNAGIKREVIKHGEFGFVVDGDCSDAIVRLLE
DKVLNKLQKNMLADIARFSGELITDQVEYYVQKMLNEHNSKGE"
complement(10167..10484)
/gene="WM1349"
complement(10167..10484)
/gene="WM1349"
/Note="ORFID:WM1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMEMOREVDEYIGQPKTGYFSPLANLARLTVEGELAREIN
HTYGEKKKXDESEANTIKAEIGDNLFVLLCLANSNGIDMTESFNETMEKFNTRDKNRF
ERK"
complement(10820..11518)
/gene="WM1350"
complement(10820..11518)
/gene="WM1350"
/Note="ORFID:WM1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIEFTMIYFVILMLPLMAOHKYSNVEKYSOVRSTSGKTR
EVALEILANGIDYDEVYKGBGFLTDHDPKKYVLSLPANDRPSVATGATIAHEVG
HAIDHOGYWFLEPRFALVPAVNLGSSLSYIMIMIGIITLTAIGSAFGSTALWIGAGLM
```

Query Match 100.0%; Score 387; DB 1; Length 333750;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 4
AP003362
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1
Kurida,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
21311952
11418146
2 (bases 1 to 346900)
Ohta,T.
Direct Submision
Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
of Medical Technology and Nursing, Department of Medical
Technology, 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan
(E-mail:tohta@tsukuba.ac.jp, Tel:81-298-53-3454,
Fax:81-298-53-3454)
On May 29, 2001 this sequence version replaced gi:13875626.
LOCATION/Qualifiers
1..346900
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
12..263
/gene="SAV1313"
13..263
/gene="SAV1313"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57475.1"
/db_xref="GI:14247084"
/translation="MLTLIKLDEQVYIYETIPEDISNGKGSYFNKKDAEVIDFS
LSEIENEYFMYLRNKSFSVVRDFLEKQEPENYKIAMY"
987..1181
/gene="SAV1314"
987..1181
/gene="SAV1314"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57476.1"
/db_xref="GI:14247085"
/translation="MMLQSYIGIKLVKLTLNKKILIGKVIDPDKVDNPDGNSIEID
TGRITIDISENKIKITVLRK"
complement(1725..2750)
/gene="SAV1315"
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
```

```

gene
CDS
    /translation="MISFENDYIEGAHEVYLRKLVDTNMLVOASGYGDFDQFTQAIEKI
    KDTIDCPNATIRFLVGGTQTNVYINSMLESEGVISADTGHVAHGAIEYSGKRV
    ITIPSEKGVASADSEYTMETFEKSDKDMWEPGWYIISHPTETGLYSKSELEIC
    KYKQYQLDFLMFGARLGYLMSDOSDITIKDIAYKCDVEYIGGTGIGALCGSAIYFT
    KNNEPQFTTRIKHHGALLKGRLLTGIOEMLETTDMLYFNISRHAIEMNKMKGDKFN
    KGRLYFDSPTNQOFFILSNEKIAELQKVKFAMVKYDDQHHVYRFAISMATTEENL
    NKULELI"
    /gene=".3217"
    /gene="SAV1316"
    /gene="SAV1316"
    /gene="SAV1316"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAB5748.1"
    /db_xref="GI:14247087"
    /translation="MEQIKLKTFTATLELLEKNINAPLSSEATNMLKVNITIKETI
    ERFPNNEEFNALILTSYK"
    /gene=".4755"
    /gene="SAV1317"
    /gene="SAV1317"
    /gene="SAV1317"
    /codon_start=1
    /transl_table=11
    /product="cardiolipin synthetase homolog"
    /protein_id="BAB57479.1"
    /db_xref="GI:14247088"
    /translation="MEFVPEFIEGKTMRYTFSNDLGTFTIILAIGFIINLVLAFTII
    FLERNRTASTWAMLEVLVPLIGFTIYLFPGRTVASRLKNNGNVLTDDGILK
    OQIESPDKNGYDGNKQKQKHHDVLSMLDODGFLTENNKVDHFTDGDNDLYQVYKLD
    IKNAKEYIHLEYTPRALDGLGKRITLALBEEKLQGLQEVKLIDVDGSKNKKMNEF
    KSLGGEVEAFASKLPDLNFRNNRNHRIITVDGQLYGVGGNIDEXVLGKRLIYW
    RDTHLRIQDADVALQLRFLIDMNSQAHNRQPEYDVYKPKNGPLGNSPIQIAAGCP
    ASWMOIEGYGYKIMSAKSYVLOSPIYIPDINSYINAIKIAKSGVDHAIIMPCPD
    HPLVYATFENASDLSSGVKIYTYENGFTLSKMLIDDEIYVSGTANDERSEFLNF
    EVNAFYVDENLAKDLVAAYEHDITKQQLKESYANRPLSVKFKESLAKLYSPYL"
    /gene=".5436"
    /gene="SAV1318"
    /gene="SAV1318"
    /gene="SAV1318"
    /note="ATP-binding protein"
    /codon_start=1
    /transl_table=11
    /product="ABC transporter homolog"
    /protein_id="BAB57480.1"
    /db_xref="GI:14247089"
    /translation="MIOISNINKSFNKRGVLNKISPDIEGKCIATLIGKNGACKSTLI
    DLIGVNAANSGEIFPDKDLQSENSIMOKRMPDQKVEIITNLXQSPVENPPL
    ERTIELTKDSSQLNQFNKLSGGQRLDPLVSLIGPQLILDEPTSMQIEIREY
    PMSIT"
    /gene=".5838"
    /gene="SAV1319"
    /gene="SAV1319"
    /gene="SAV1319"
    /note="ATP-binding protein"
    /codon_start=1
    /transl_table=11
    /product="ABC transporter homolog"
    /protein_id="BAB57481.1"
    /db_xref="GI:14247090"
    /translation="MITYSHYIEFVRMSDKIILINGEIIILNDSTSHIRTNQOSIT
    LSEDTIRKLKLDLVLQKNHGTIKITTSNVNPTILYLOQLHMLDIEIOKVSIY
    DSYFNKKQKRGSDYDKLLENRI"
    /gene=".6538"
    /gene="SAV1320"
    /gene="SAV1320"
    /gene="SAV1320"
    /note="similar to ABC transporter integral membrane
    protein"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAB57482.1"

```

```

gene
CDS
    /db_xref="GI:14247091"
    /translation="MILSYLKIEFKVIMRKKTLLILSLFPVIFYILFTSILELPEDV
    KPFVEYMSMTVSYLSLSFLTFPLDIIINEKONBMRQRLMWTPTFTSYISKYVK
    TMQAFALITIVMVGHFYKGVMSAVOMLSCIFILMLGASLITIGILFSLNDIOK
    TSLANIVITIGVAGLMPRIPTPNMLQHVAVHVPYHLRLSLGDIASNHHNLIS
    FAILLYVAGSITIAVYCISFKRAE"
    /gene=".7633"
    /gene="SAV1321"
    /gene="SAV1321"
    /gene="SAV1321"
    /note="similar to two-component sensor histidine kinase"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAB57483.1"
    /db_xref="GI:14247092"
    /translation="MKFLKDTSAEISLITLYLFPPIAGIFPNVYGPKMXYIIVF
    SLSYLIVYNNRLNLTMLPILILIHRTICFVESHPLSLFFYSAPAFPTKRN
    NVKRTANLEFILMTICTYITLYLYNNYFVAMVYVYISLMDNFKMKKREYQKE
    IAKNRHINTLAEQERHRIGDLHTGLHVFASLSLKSBLAYKLIDTVEKAKBL
    AINKLSRESLNKREIIDDVKLPSTEEIDISIRKVLKADIDFTFENKELACVSPK
    OSMVMITREAINNVVIRHANASKVHGKLTNNHKLMLMIEDDGKIDSDCEKVISIO
    RVQHLNGTIAVDSTNGTKIIIEISTGIA"
    /gene=".8232"
    /gene="SAV1322"
    /gene="SAV1322"
    /gene="SAV1322"
    /note="similar to two-component response regulator"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAB57484.1"
    /db_xref="GI:14247093"
    /translation="MTSLIAEQNNMLRQAMVOLIKHGDPEIADTDNGLDAMKLE
    EYNPVNIYDIEMPGMGLVLAIEIRKHLINIKVITVTFKRRGFEKAVVNDVAYV
    IAKRSIEBELVETRNKYNKEKESATLMSFPYDKRPLPKQIYVIREIGNGLSKEI
    SEKFLITDGVTRNVTYSIIDKLPADNRFDAMKANRKGWI"
    /gene=".8533"
    /gene="SAV1323"
    /gene="SAV1323"
    /gene="SAV1323"
    /note="ATP-binding protein"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAB57485.1"
    /db_xref="GI:14247094"
    /translation="MNISNENVTLLAVIIAIIIGIFLQIFFKPLIVTAVSILLGI
    FVGFYVLYVSFNNKRN"
    /gene=".9205"
    /gene="nuc"
    /gene="nuc"
    /gene="nuc"
    /note="SAV1324"
    /codon_start=1
    /transl_table=11
    /product="thermonuclease"
    /protein_id="BAB57486.1"
    /db_xref="GI:14247095"

```

Query Match 100.0%; Score 387; DB 1; Length 346900;
 Best Local Similarity 100.0%; Pred. No. 1,2e-77;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GATCTTCCTCCCTTCACCAAAATGAGAAACACGTCGATCTACAGTCCACGAAGACTA 60
    |||
Db 200303 GATCTTCCTCCCTTCACCAAAATGAGAAACACGTCGATCTACAGTCCACGAAGACTA 200362
    |||
QY 61 AACCATGTGACCCGTGATATGAGATACGGTTACACCAATCCTAATGATAGAAATCATATACA 120
    |||
Db 200363 AACCATGTGACCCGTGATATGAGATACGGTTACACCAATCCTAATGATAGAAATCATATACA 200422
    |||
QY 121 CGTCTGACCATTTCCATATTAATCTACTTGTGTAACCGCTAATACGACCGGTTTTTTAG 180
    |||

```


Db 200423 CGTGTGTAAGCATTTCCATTAATTAATCTACTTTGTTAAACCGCTAAATACGACGGTTTTTTAG 200482
QY 181 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 240
Db 200483 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 200542
QY 241 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 300
Db 200543 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 200602
QY 301 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 360
Db 200603 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 200662
QY 361 TTAAACATTCACCTGAGAAATAAATAC 387
Db 200663 TTAAACATTCACCTGAGAAATAAATAC 200689

RESULT 5
AF270032 3269 bp DNA linear BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1042f07 genomic
DEFINITION sequence.
ACCESSION AF270032
VERSION AF270032.1 GI:3623936
KEYWORDS
SOURCE Staphylococcus epidermidis
ORGANISM Staphylococcus epidermidis
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 3269)
AUTHORS Kimerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torturella-Miller,I.,
Listenebee,S., Ashanti,C., Alshuler,G., Mamo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3269)
AUTHORS Taylor,J.David., Kimerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torturella-Miller,I.,
Listenebee,S., Ashanti,C., Alshuler,G., Mamo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source 1..3269
/organism="Staphylococcus epidermidis"
/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1042f07"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN

Query Match 67.8%; Score 262.2; DB 1; Length 3269;
Best Local Similarity 79.8%; Pred. No. 3.1e-49;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Db 2417 TATCATTAAGCATTTCAAGATTATCAACTTTTATTCACAGCTAATACACAGGTTCTTAG 2476
QY 181 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 240
Db 2477 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 2536
QY 241 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 300
Db 2537 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 2596
QY 301 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 360
Db 2597 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 2656
QY 361 TTAAACATTCACCTGAGAAATAAATAC 387
Db 2657 TTAAACATTCACCTGAGAAATAAATAC 2683

RESULT 6
AX145350 3269 bp DNA linear PAT 31-MAY-2001
LOCUS AX145350
DEFINITION Sequence 4072 from Patent WO0134809.
ACCESSION AX145350
VERSION AX145350.1 GI:14283915
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Kimerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4072 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source 1..3269
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN

Query Match 67.8%; Score 262.2; DB 6; Length 3269;
Best Local Similarity 79.8%; Pred. No. 3.1e-49;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1 GATCTTCTCTCTCTTCACCAAAATGAGAAACAATGCAATCTAATCAAGTACCAAGACCTA 60
Db 2297 GATCTTCTCTCTCTTCACCAAAATGAGAAACAATGCAATCTAATCAAGTACCAAGTACCTA 2356
QY 61 AACCATGTGACCTGATATCGATACGATACGTTCCACCAAAATCTAATGAATGAATCATACA 120
Db 2357 ATCCATGTGACCAAAATGAATGATATGATATGATATGATATGATATGATATGATATGAT 2416
QY 121 CGTGTGTAAGCATTTCCATTAATTAATCTACTTTGTTAAACCGCTAATACGACGGTTTTTTAG 180
Db 2417 TATCATTAAGCATTTCCAGATTATCAACTTTTATTCACAGCTAATACAGGTTCTTAG 2476
QY 181 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 240
Db 2477 ATTGTATATAAATTTGAGCGACATTTCATCGCTTTGTGTAATCCTTCAGCAGCTTAA 2536
QY 241 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 300
Db 2537 CCATATAAATAATTAACATCCGCTTCATCTATGAGGATTTTCGCTGCGCTCTAATTTGTG 2596
QY 301 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 360
Db 2597 TTTGGAATGTCATCACCATTTCATATACCACTGTATCAATTAATTAATGAATCATGTG 2656
QY 361 TTAAACATTCACCTGAGAAATAAATAC 387
Db 2657 TTAAACATTCACCTGAGAAATAAATAC 2683

Db 2657 TTACCACTCGCCAGATGAATAATAC 2683

RESULT 7
AE016747
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AE016747 300892 bp DNA linear BCF 01-Jan-2003
Staphylococcus epidermidis ATCC 12228, section 4 of 9 of the
complete genome.
AE016747 AE015929
AE016747.1 GI:27315369

Staphylococcus epidermidis ATCC 12228
Staphylococcus epidermidis ATCC 12228
Bacteria: Firmicutes; Bacillales; Staphylococcus.
1 (bases 1 to 300892)
Zhang, Y., Ren, S., Li, H., Yu, G., Lu, L., Lu, G., Jia, J., Tu, Y.,
Qin, Z., Chen, Z. and Wen, Y.
Direct Submission
Submitted (05-NOV-2002) Chinese National Human Genome Center at
Shanghai, 250 Bi Bo Road, Shanghai 201203, China
Location/Qualifiers
1..300892

/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:116280"
113..346
/locus_tag="SE0907"
113..346
/locus_tag="SE0907"
/codon_start=1
/transl_table=11
/product="acyl carrier protein"
/protein_id="AA004504.1"
/db_xref="GI:27315370"
/translation="MENPKVVDIYDRGVDADKYTEASFKDDLGASDLDAELVM
ELEDFGTIPDEAEKINTVDAAVINSLEK"
461..1198
/locus_tag="SE0908"
461..1198
/locus_tag="SE0908"
/codon_start=1
/transl_table=11
/product="Rnase III"
/protein_id="AA004505.1"
/db_xref="GI:27315371"
/translation="MANOKKKEVHNFOOKFTDDMKSLGIHFKNIDLYOAAFSHSSI
NDENMRLEHNERLEFDGAVLEIYSRYLPRRPHLPGLSGNLTKMATIYCEPSLYIF
ANKIKNELILGKGEKTRPSLISDAFEAFALYLDGGDLQSVTFPAKVIAP
IYVDELVGVDVFKTOFOEYVHSQNGDYDYLKEEGPAHRLFTSEVILEKAAVE
GKGTKESEQAEEQAVKLMKNKSL"
1325..4894
/locus_tag="SE0909"
1325..4894
/locus_tag="SE0909"
/codon_start=1
/transl_table=11
/product="chromosome segregation SMC protein"
/protein_id="AA004506.1"
/db_xref="GI:27315372"

/translation="MYLKSIDALIGKSPADHTNVQDKGVAIYGPNGSGKSNITDA
IKWVLEQSAKSLRGAKMEDIIFGSAEHRKAQNAFAVKILDNHSOKLIDISELVYT
RLRSGSESEYVLYNNRARKLIDILFDSLGKAEAFSIISGGRVDEILNARKINRQ
IIEESAGVLYKKRKAESIOKLDHTEEDNLRVADIIYDEGRVPELKEEAAIKRKO
LSKEMOSDVITVSDIDHYTEDNORLDRKLSQAEEKGEOAOIOLLOKYGK
ROONDYDIKLVLEYKATENYEQLSKLVNLEKRNKOSERTNAREEEDLNIESID
STIKERAKNEKLAEILKNOKOLNKEVDESLIYISDQHDEKLEIKNSYTTLASE
QSDVNDIDFLEHTINENAKSRSLDSRIVEAFNQDKIDIONITOTQKEVSSKSGME
KYEONIQOOLEQTLDSKRLSEYENKLYOARYNEKLSRIDSILATOEEDYTYFFNGV
KILKAKDELRIGHGAVAEVINPSEMTQATELTAGASLOHYIYDNEKROGAIOYL
KORGCRATEFLPLNVOPRHVADIKDVAKSGCFNTINASDAIYNVSAKYONITENLIG
NTIYENLKHANELARAIYRIRIYITLEBDVYVPGGSMTGGAARKTKSILSODELST

MRNQLDYQRTAFEEFORKEQTOAEQISEQYFSAQOYNINKEQVHHHELELRLK
TDEAHLKNEHEEFERKNDGYQDSKSEYTLKEXONLIEIQOOLKQLESDIENTEYOLS
KEGKASTHQOOLHOKOSDLAVAKERIKSOQOYERLDKOLSDSEROKIENVEKIKL
FNSDEMGKDAFEKLEKQIOQOENVRQNLNQSLSEIKQORLDNERIENESOLQKCH
ODLSTENHNOFKAKOSKLDVLIINAHIDHNLNTYOLTYERABAREYDSEPTDNLKPK
VKLTKEISIDELKRYFNALIEQELNERYTEFLNEQPTDIREAKETLEPQIKHENDKEVE
GRKTTFHAVQDHFTTVFKOLREGGQAEILRTEDDILSKGVDIYVPPCKQLQHSLS
SGERALSAILALLFALKVRSAPEVILDEEALDEADANVIRIAQYINLSTETQPTVI
THRKGTFFESDRLYGVYTMQESGVSKLSVNLNTIDVMEQEA"
4891..6117
/locus_tag="SE0910"
4891..6117
/locus_tag="SE0910"
/codon_start=1
/transl_table=11
/product="signal recognition particle"
/protein_id="AA004507.1"
/db_xref="GI:27315373"
/translation="MSFFKRLKDESSKSEDDIQKLDLETVDNSNANSDSQSDPNNSD
EQVPRPKPKLSLSEADFDGDLISDFEETIEAQIGAKFAGLERSKONFOQLNLN
IARYRVDEDFEALREMLITADVGFTVWKLDELRTAEORNIQETDRLREYIAEK
IVEIYHOEDHSEAMNIEDGRLVNIVMGVNGVYKTTTIGKLARYQOEGKKVYLAAG
DFPRAGAIQOOLNMGVRGVEYVSONEGSDPAVVYDANINAAKNGVDILICTAGRL
OKNSINMORLEDKMKRVINRAIPDAPRPAELICDATTGONALSOARBSFETVNGSVL
TKLDGTAKGCTVLAIRNELHIPKYVGLGKMDLQPFNPESIVYGLPDMIQNNDI
PEEISRNSSVSEEGN"
6119..6451
/locus_tag="SE0911"
6119..6451
/locus_tag="SE0911"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA004508.1"
/db_xref="GI:27315374"
/translation="MGONDLVYKLRMNYLFDEYOSLITNKOKYLFEPFLQDYSLSI
ADTFEYRQAVYVNIIRGTDLVYDEYSKLRLLYORFEXRRELYILMQSLNPPELLKQY
ITOLEELE"
6482..7849
/locus_tag="SE0912"
6482..7849
/locus_tag="SE0912"
/codon_start=1
/transl_table=11
/product="fih protein"
/protein_id="AA004509.1"
/db_xref="GI:27315375"
/translation="MAFEGISDLRLQATPMOKMRGKGVTEADIKTMREYRLALLEADV
NFRVYVEPKVTVSEKALSGSYMOSLTPGOOYIYIYQBELTSLMGENTSTKMNKRPPT
VWVWVLDQGGKTTTAGKALMLKRRKYNKPLVAGDITRPALIDOLQTYGKQIDIPY
YVSGDIVSPQOIVENALHAKKEHLDGVLIIDTAGDTRIGDALNNEQVEIKSPDEI
MLVVDAMTGODAVNAQSPDEQDVSGVLTJTKLDGDTIGDALNNEQVEIKSPDEI
SEKLDLELELHPRMASRIIGMGVDVLTLEKAAQVDQAEAKLEKMKESSFTLDFE
LEQLDQVKNLNGPLDIIMKMPGNNKKKGLDKLMSKQIDHIKAIIOISMTPKRSNDE
TILVSRKRIKAKSGSLQEVNRLMKQFNDMKKMMQFTCGCGKAKAKRNQNMNLKG
MNLFP"
8281..8556
/locus_tag="SE0913"
8281..8556
/locus_tag="SE0913"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S16"
/protein_id="AA004510.1"
/db_xref="GI:27315376"
/translation="MAVKIRLITRLGSKRNPFRIVVADARSPRDGRLIQLYGNPAH
VNAPEVKIDBELAKMLHDCAKFTDYVHNILSNREGILKKFDEQKNK"
8683..9186
/locus_tag="SE0914"
8683..9186
/locus_tag="SE0914"
/codon_start=1
/transl_table=11

/product="rimm protein 1"
 /protein_id="AA004511.1"
 /db_xref="GI:27315377"
 /translation="MEVEGLQVNHGICGEVKAIVKNSDFTETRFQEQGLVKNHT
 EIVYVASVRIHKGHEFQENINNTNELHLKGDVYDERHQDIELGHEYYSDI
 ICGYFQRKDDPPIGVRINIFETGANDVWVVKGEKELIPIEDVVDIDIEFTIKIT
 PMEGELD"
 9186..9923
 /locus_tag="SE0915"
 /locus_9923
 /locus_tag="SE0915"
 /codon_start=1
 /transl_table=11
 /product="tRNA (guanine-N1)-methyltransferase"
 /protein_id="AA004512.1"
 /db_xref="GI:27315378"
 /translation="KIDYTLFPEMEEGVLAHNSILKRAODKIIINVNTINFDYSIN
 KHNQYDVFEGGQGVNLKPEVFNAMEDINNEHRTVIRLPCOGREPFQELAOEISEE
 AKHIVYIGCHEYGEYERIKHLVTEDEISMGDVLTGGELEPAMTMAIVLIGVGLN
 OASHDEQSFSDLEEFPEYTRPREYKNMSVPEVLLSGNNAHIDQMRHEKLIPTYEKK
 PDLQYVFLTEKDELDLEETYYKKKLKND"
 10029..10379
 /locus_tag="SE0916"
 /locus_10029..10379
 /locus_tag="SE0916"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal protein L19"
 /protein_id="AA004513.1"
 /db_xref="GI:27315379"
 /translation="MSNHLKVAVTKSQLRTDLPFRSGDTLRVHVAIVGSESRIDQV
 EGVGVYIKRGGGISTEFTYVRKISSGVGERTPELTPPKIEKIEVKRKGAAKRLYYL
 RSLRGKAAIOEIR"
 complement(10910..13513)
 /locus_tag="SE0917"
 complement(10910..13513)
 /locus_tag="SE0917"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AA004514.1"
 /db_xref="GI:27315380"
 /translation="MIKNIWMPKIRFTLIVASCLISFLVYEPILKQPTITKGVSG
 SGGDROMPOMHLXENHTSKGIFDASFGDGVLSAIVYSPLMNIINSMTM
 LLEQITVNPPIHSWTINQILMAVRYVITFESLYSLIKAPMPTAIILYGM
 SVIVYVYETFAEYCNLILMPMSLIAIERFEKERIGWIFPAIVTLFENFYSYE
 AIVIEYVYTRALAIPEKIOVMOKLILYLCAVTLISLVASIGLTVGSSPLDNDST
 ONPKRITFETLPEENYIPADGFIITSLAIILAFCKRLQHYHYKFAITATVIL
 LIGSSQNDASFLNENSPIDRRWVYLAISTVLALPFLQDSEISIKETFPVAPVE
 IYGFIFLALSSESVKMPALILILYLFIRKSLILTRTSMVLAIVLELAQVLEMT
 NDNRITILEPPIYITKTINDSSYRSFVLPKKIKRYMQSSDPLKRDYFSYVALNSPL
 IYHVGTSLYSSIPGDLILKXYDQTLQIMPYDKNSTYRLNNRANIMSLMDVDRIL
 HPDDLNNPGEKFKELITDKKQIMINSVNTINPASHITRKIVDARKLGLPDRQGM
 LKGVVLNKRSGQANTQPKPAPNLSLNKQNLNNANMIDSKILKVOHNGVTLNLRPI
 VKNYDGVYREKMGVILGEGYQTLKAKSQDQPKYKERNRNGPFIKKKRGVILPDM
 IYAESEHAMANAKRPLKVOOGNGIMTITPIYEGOTKIKLSTTPPYFILLTSCIGIIL
 SILTFHYVYKRR"
 complement(13506..15779)

Query Match	Best Local Similarity	Score	DB 1:	Length
Matches 309; Conservative	79.8%;	262.2;	300892;	
		Pred. No. 26-49;		
		Matches 78;		
		Indels 0;		
		Gaps 0;		
QY	1	GATCTTCCTCCCTACCAAAATGAGAAACATGCACTTAACAAAGTCACCAAGACCTA	60	
Db	296677	GATCTTCGATTCCTTTATTTAAAGTTTCAACACATGCACTTGCAAAATTCCAAGTCCTA	296736	
QY	61	AACCATGTGACCCGTGATATCGGATCGGTTCCACAAATCCTATGATGAATCATCTACA	120	
Db	296737	ATCCATGTGAACACGAAATAGATATGTGATCTCCAAAGCCTTAAGATATGAAATCATACA	296796	
QY	121	CGTCTGACGACATTTCCATATTTATCTACTTTGTAAACGGCTAATAGACCGGTTTTTAG	180	

Db	Accession	Source	Organism	Reference	Authors
Db	296797	TATCATTTACCGCATTTTCAAGATTATATCACTTATTTACAGACTTAATACACAGTTTCTTAC	296856		
QY	181	ATTGTGTATAAATTGAGCGACCAATTTTCATCGCTTGTGTGTCAAATCCTTCACGACGCTTAA	240		
Db	296857	ATTATATAAGCATTTGTGTGCGACCAATTTGCTACTTGTGTAGTGTCTTCTTAACATTTGA	296918		
QY	241	CCATTAATAAATTAATACATCGCTTCATGTATGAGCGATTTCTGCTGCTGCTCTAATTTGTC	300		
Db	296917	CCATTAATAAATATGATGATGCTATGCTTTCTTCAATTTGCTATTTTCTGCTGTGACGAGATTTGAG	296978		
QY	301	TTTGGAATGATGTCATCACCACCAATTTTCAATACACCGCTGTATCAATATATTTGAATCATGTG	360		
Db	296977	TTTGGAATGATGTCATCACCACCAATTTTCAATACACCGCTGTATCAATATATTTGAATCATGTG	297037		
QY	361	TTTAAACATTTACACCTGATCAATTAATAC	387		
Db	297037	TTTAAACATTTACACCTGATCAATTAATAC	297063		
RESULT 8					
LOCUS	AE017028/c				
DEFINITION	AE017028	290117 bp	DNA	linear	BCT 30-Apr-2003
ACCESSION	AE017028				
VERSION	AE017028.1	GI:30255149			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					

gene		EAKFNMRGKGKDIKTDRBGRIT" 7854. .9680 /gene="pepf-1"	
Query Match	54.4%;	Score 210.6;	DB 1; Length 290117;
Best Local Similarity	71.7%;	Pred. No. 9.2e-38;	
Matches 276;	Conservative 0;	Mismatches 109;	Indels 0; Gaps 0;
Oy	3	TCCTCTCCCTCTCACCAGAAATGAGAAACATGCATCTTAACGAAGTACCAAGACCTTAA	62
Db	280879	TCCTCTCAATCTTTTGGAAATCTCTGACGCTTCATATAGTACCTAATCTTAAT	280820
Oy	63	CCATGTACCCCTGATATCGGATACGGTTACCAAAATCTTAATGAATGAAATCATACAG	122
Db	280819	CCGTGTACCTAGATGAGATGGGATGGTCCCAAAATCTTAATGAAATCATATATG	280760
Oy	123	TCGTACGATTTTCATATATCTACTTTGTTAAACCGCTAATACGACCGGTTTTAGAT	182
Db	280759	TCACGTGGCATTTCCGATGTCCACTTATTTACTGCAAGTACAACTGGTTATTATGAA	280700
Oy	183	TTGTATTAATTTGAGGACATTTGATCGCTTGTGTCATACCTTACGACGCTTAC	242
Db	280699	CGATATATATTTTGTGAACTTCTTCACTCCGCTGCTGTACACCGTCAGACCATTTGTC	280640
Oy	243	ATAAATAATATACATCCGCTTCATCTATGCGCATTTCTGCTGCGCTTAATTTGTT	302
Db	280639	ATAAATAATATATACATCTGTTTCATGCTACTGCTACCTGTTGACGATTTGCTC	280580
Oy	303	TGGAATGGTGCATACCAATTTTCAATACCACTGTATCAATATATTTGAAATCATGTG	362
Db	280579	AAAAAATGGCTGCTGCTCAATATCAATCCACCTGTATCAATATATTTGAAATCATGTG	280520
Oy	363	AACCATTCACCTGAGAAATTAATAC 387	
Db	280519	AACCATTCCTCAGCGCTAATAATAC 280495	
RESULT 9			
AE017002/c	304680 bp	DNA	linear BCT 16-MAY-2003
LOCUS	AE017002	304680 bp	DNA linear BCT 16-MAY-2003
DEFINITION	Bacillus cereus ATCC 14579 section 5 of 18 of the complete genome.		
ACCESSION	AE017002	AE016877	
VERSION	AE017002.1	GI:29894935	
KEYWORDS			
SOURCE	Bacillus cereus ATCC 14579		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
REFERENCE	1 (bases 1 to 304680)		
AUTHORS	Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fongstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.		
TITLE	Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis		
JOURNAL	Nature 423 (6935), 87-91 (2003)		
MEDLINE	22608415		
PUBMED	12721630		
REFERENCE	2 (bases 1 to 304680)		
AUTHORS	Candelson,B., Galloux,K., Ehrlich,D.S. and Sorokin,A.		
TITLE	The number of ribosomal RNA operons in Bacillus cereus		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 304680)		
AUTHORS	Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fongstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France		
FEATURES	Location/Qualifiers		

source	1. .304680
organism="Bacillus cereus ATCC 14579"	
/mol_type="genomic DNA"	
/strain="ATCC 14579"	
/db_xref="ATCC:14579"	
/db_xref="taxon:226900"	
68. .427	
/locus_tag="BC1239"	
68. .427	
/locus_tag="BC1239"	
/locus_tag="BC1239"	
/codon_start=1	
/transl_table=1	
/product="hypothetical Membrane Spanning Protein"	
/protein_id="AAP08224.1"	
/db_xref="GI:29894936"	
/translation="MLRRKLLYFLVPLPYLAWLISMTKIELMALFLGVPLFESLNRI QEOSILEICLFSISLEFSTVSIVLLELFLRHSFTLMKFWNITLDAICTVYFVVL KEIVGKQVFEQDNKKE"	
551. .2041	
/locus_tag="BC1240"	
551. .2041	
/locus_tag="BC1240"	
/codon_start=1	
/transl_table=1	
/product="L-lactate permease"	
/protein_id="AAP08225.1"	
/db_xref="GI:29894937"	
/translation="MALIPIMMIFICLFKQTSLRSSLISVAVSGVYLLSPFOL GISEVHATIKGWLICFIVGVLEFGIFLHLMKKQYIDQVARELEEVQDRLQWL LMGICGPLLESVSGFICFMVAPDIFSLGVRPQVLLSFGLASSWGAMATGRI IGSQLNMPPLTTGNTALLSIPFAVILLVGVGMQAVIEKMEGVFLFEL GYLSNMPVSEVLAGILSSIVTIFGFLIKLKGSGNLIIEHAATREVSIIKII SPYELFVCLISLRVPALDPRSVAVDLKSYSTELYSGLPMLGMPCECTIIE FRIPSNIRKOSLSQTIKQWIPATTTMFLAISLMSGMSHLATADTDETFPV FVAPFICGIFLGSNAGSNMFTLQMDQVAVLPMQVITLQMTASVATIIACP SRIYLGAVLNICNIPYRENELKTTLMFGAVLVVEIVWYMLRN"	
2132. .2818	
/locus_tag="BC1241"	
2132. .2818	
/locus_tag="BC1241"	
/codon_start=1	
/transl_table=1	
/product="hypothetical protein"	
/protein_id="AAP08226.1"	
/db_xref="GI:29894938"	
/translation="MTLKNYGVLKGYIOSEIGCKTPHYOVHLOGETGVDRYAINV KSOSEPSVFLYFASNNIKSEAIHILPPLPGFTEIKNNRPVALDYIRGNLFDSNOMI PLPEKAGADNDLNEKITERYIKRAIEKAIIVAGERGGEENTPDSYFHKPONGIH DIHMOGNAGVEMQDGDGIMQDGILIFHEKEEWIGIFLAFOSQSMCTDEBGRAPV EHCQYKGNIK"	
complement(2830. .3663)	
/locus_tag="BC1242"	
complement(2830. .3663)	
/locus_tag="BC1242"	
/codon_start=1	
/transl_table=1	
/product="Integral membrane protein"	
/protein_id="AAP08227.1"	
/db_xref="GI:29894939"	
/translation="MELVNSIFQILLTSVLIQSLISGVIIVIGFLGYSLETRMYS RAEFRKGFELTAWGVPIHEHAIKMCVLRHKIVAQFPPTDSOGALGVQHOYNO KSVYORGNFEGIGSPIISGITALLIMRVEVPESYFLPNTLEKTIASISINTMIQ NMULTSVLLKSLFTTINLNPSPMLFLFALICSAHIAASKPIKSIQSVIYMFIV LELFNITGLFQYDSNOLIGOVKTYNMLIAFSVALFSCISTLVSPFKIRGNS F"	
3860. .4630	
/locus_tag="BC1243"	
3860. .4630	
/locus_tag="BC1243"	
/codon_start=1	
/transl_table=1	
/product="hypothetical protein"	
/protein_id="AAP08228.1"	

Query Match	Best local Similarity	Score	DB	Length
Matches 275; Conservative	71.4%; Pred. No. 2,1e-37;	209;	DB 1;	304680;
	0; Mismatches 110; Indels 0; Gaps 0;			
QY 3	TCCTCTCTCTCTCTACCAAAATGAGAAACACGTGATCTAACACGTACCAAGACCTTAA	62		
Db 244523	TCCTCTCTCAATCTTTTGAAATAATGTTGTGACGCTTGTCTAGTAAGTCACTAACCCAT	24446		
QY 63	CCATGTGACCCGTGATATCGGATACGGTTCCACCAATCTCTAATGATTAATAATCATACAG	122		
Db 244463	CCGTGTGTCTCGAAATTTGGGAATGGTGTGCGCAAACTTAATGCTTAAAGCTCAATTAATG	24440		
QY 123	TCGTGTGCGATTTCCATATTTATCTACTTGTGTTAACCGCTAATACGACCGGTTTGTAGAT	182		
Db 244403	TCATCTGCGATTTCCGGAATGTCTATTATTACTGCAAGTACACCGGTTTATTGAA	24434		
QY 183	TGTATATAATTTTGACGACCATTTCAATCGCTTGTGCAATCTTTCACGACGCTTAAC	242		
Db 244343	CGATATAAATTTTGTGACACTCTTCTCATCAGCTGCTGTACACCGTACGACCGTTGTC	24428		
QY 243	ATPAAAAATTAATAACATCCGCTTCATCTATGGGATTTCTGCGTGCCTTAATTTGGT	302		
Db 244283	ATPAAAAATATATACATCCGCTTCATCTAATGCTACTTACGCTGTGACGAATTTGGT	24422		
QY 303	TGGAATGTGATCAACCAATTTCAATACACCTGATCAATAATTAATAATCATGTGT	362		
Db 244223	AAAAACGCTGCTGCTTCATTAATTTCCGCCCGATCAATATATTTAAATTCATGATTT	24416		
QY 363	AACCATTCACGTGAAGATAATATAC	387		
Db 244163	AACCATTCCTCCAGCGCTATAATATAC	244139		
RESULT 10				
LOCUS	AL596170	313450 bp	DNA	linear
ACCESSION	AL596170	AL596170	complete genome, segment 8/12.	
VERSION	AL596170.1	GI:16414292		
KEYWORDS				
SOURCE	Listeria innocua			
ORGANISM	Listeria innocua			
REFERENCE				
AUTHORS	Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baguerio, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Chabild, A., Chetoui, F., Couve, E., de Darvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussange, O., Entian, K. D., Fsihi, H., Portillo, F. G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kref, J., Kuhn, M., Kunst, F., Kurpik, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E.,			

Kurapkab,G., Madueno,E., Maltournam,A., Vicente,J.M., Ng,E., Nedjari,H., Nordsiek,G., Novella,S., de Pablo,B., Perez-Diaz,J.C., Purcell,R., Remmel,B., Rose,M., Schluter,T., Simoes,N., Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P. Comparative genomics of *Listeria species* Science 294 (3543), 849-852 (2001)

JOURNAL
MEDLINE
PUBMED
11679669
2 (bases 1 to 347050)

REFERENCE
AUTHORS
TITLE
JOURNAL
Glaser,P., Frangeul,L. and Rushtok,C.
Direct Submission
Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96. Fax: +33 (0)1 45 68 87 86.

COMMENT
FEATURES
source
Location/Qualifiers
1. .347050
/organism="Listeria monocytogenes"
/mol_type="genomic DNA"
/strain="EGD-e"
/db_xref="taxon:1639"
complement(95. .841)
/gene="lmo1688"
complement(95. .841)
/gene="lmo1688"
/note="similar to glucose 1-dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="CAC99766.1"
/db_xref="GI:16411142"
/db_xref="SPTREMBL:O8Y6J6"
/translation="MNKVALTVSSSGKGRGRTAIALAKGYDIANVSPRRKAEPO
OEIDOLGRVCYEFKANGDVEKREKRVDEFEGRDIFINNAAGVRLPMELEES
HWDWTMINAKALTRGDEPAKLMQHOSGKITISSTGSIYRLLEYTVGVSKAAVE
SLTRYLAELAPFGIAVANVSGGLIETDALNHPNREELKDAVSKTPAGRIEPLDL
VNAVFLASEKMDMIRGOTILVDGGRTLLV"
complement(join(845. .1942,1948. .1953))
/gene="lmo1689"
complement(845. .1942)
/gene="lmo1689"
/note="similar to A/G-specific adenine glycosylase"
/codon_start=1
/transl_table=11
/protein_id="CAC99767.1"
/db_xref="GI:16411143"
/db_xref="SPTREMBL:O8Y6J5"
/translation="MRKKRLTWDETKITAFQDALVSWYEANKRVLPMRENTPEYRIW
VSEIMLOQTQVTVIPFNEMTOPPTMESFVNADADILKAMEGLGYSRVRLQTA
MKQVADSGEVPDITLISLKGVPYAGALISAYNQAEPAVDGNMRYIARLE
IGEDIMKASTRKIFEEVLVOLDKKNPAFNOGLMEIGALVCTPRKPMCIPLQPC
BAHKNGCVENYPVKIKKYMKTRELSTIVISEDGKIAIEKRPENGLANMNOFPTE
ISKRENDEYAKIQFLHNYGLEVLLEDEPFAHKHVSFLWMKMDIRVALQSAIPEN
WYFATEEEKKRLAFPPVYOKMOMAKDFEKE"
complement(850. .855)
/gene="lmo1688"
complement(850. .855)
/gene="lmo1688"
complement(1948. .1953)
/gene="lmo1689"
join(2129. .3123,3130. .3147)
/gene="lmo1690"
2129. .2134
/gene="lmo1690"
2143. .3123
/gene="lmo1690"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99768.1"
/db_xref="GI:16411144"
/db_xref="SPTREMBL:O8Y6J4"
/translation="MDTGHVVMGIALGALATVDPVVASOAAIGIMTATIIIGSIDP

DIDTVLKKNNADYIRNRGITHSLPMLAIWPLLSSILYLLFPATFTIHLMLPFA
VGHILPVDIEFNAGVAPPEKTVAFGFIEFDEFGISHVAIAAMLGSPVPT
FTLIYVILAYVAPRPOBMKHAVONLIPSEETITASTIEFQWRVAVTKKHVY
VGRAPFRNISTIEKFPRLPVDNETIRSKAKKKNLAIFSFVYWMRIEKKIDGTIV
TFPDLKRSNGHPFVAVVLDDDLKIVSSYGTWTFSTERKYLKLPVSI"
complement(join(3130. .3147,3156. .3617,3625. .3630))
/gene="lmo1691"
3130. .3147
/gene="lmo1690"
complement(3130. .3147)
/gene="lmo1691"
complement(3156. .3617)
/gene="lmo1691"
/note="similar to deoxyuridine triphosphate
nucleotidylhydrolases"
/codon_start=1
/transl_table=11
/protein_id="CAC99769.1"
/db_xref="GI:16411145"
/db_xref="SPTREMBL:O8Y6J3"
/translation="MKVRGFEVYNESRRFPPTISLPTRGDKSGAGYDFPSMETYTI
VPEKHIEFTDITKSTMOEDVELNTIVRSSITGIRKGLLNGTGIISSTYSNGNDGN
IGIAIKNFSENPVSIAGEVAVGVQKYLAVADTIVANESRVGVSIGR"
complement(3625. .3630)
/gene="lmo1691"
complement(3629. .3862)
/gene="lmo1692"
complement(3629. .3862)
/gene="lmo1692"
/codon_start=1
/transl_table=11
/protein_id="CAC99770.1"
/db_xref="GI:16411146"
/db_xref="SPTREMBL:O8Y6J2"
/translation="MTYVDELTLAELMVKNPMLDKRQALMWIEMLSDESSYAKAGY
PYRGPEYADVADYNOQIEHRGSHFYQVERDPYK"
complement(join(3864. .4673,4680. .4685))
/gene="lmo1693"
complement(3864. .4673)
/gene="lmo1693"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99771.1"
/db_xref="GI:16411147"
/db_xref="SPTREMBL:O8Y6J1"
/translation="MKTISVQOKKERYNIFIDEKYNVSDEVLARYOLMKGRVL
TEADIEIRQADMVRKGLNKAINFLSHRYRSEKEIDYRKQEMEPFADILKILAN
MYINDFEERALEYTKQITTKGPTTIRREYKGLTRETISRVEESYDEMOLENA
TKOAKIMKRNKRSKMKMOKITITDLOKGFSTELAKAATATESELDVADAEILQ
KOIEKTMKRNKRSISIAOKTITSLMOKGFSYDTIQSLTENEISFEERE"
complement(3872. .3877)
/gene="lmo1692"
complement(3872. .3877)
/gene="lmo1692"
complement(4680. .4685)
/gene="lmo1693"
join(4757. .5673,5681. .5708)
/gene="lmo1694"
4757. .4762
/gene="lmo1694"
4771. .5673
/gene="lmo1694"
/note="similar to CDP-abequose synthase"
/codon_start=1
/transl_table=11
/protein_id="CAC99772.1"
/db_xref="GI:16411148"
/db_xref="SPTREMBL:O8Y6J0"
/translation="MHILTGATGFIGDHLVHELEKSDHELYTLTROKKNANWHYI
EVLNEDKLPNLEDLPVDCINTLAGAGLMEKWTYDKKVIYNSRIEATSAALSTYKMK
KSKPKLIMASAIQATYSSKSTIYLTDEBNTVADNFIQATYVEWETASAADLQIRV
VYARFGLVGTNGSPFVEKLFQYTGFRFGNGQWYSMIHVDDVAAILTFIDHEQ


```

gene
CDS
    /codon_start=1
    /transl_table=11
    /protein_id="AAC8394.1"
    /db_xref="GI:1146197"
    /translation="MKVKLVVSMULSIAFVILMLNPFPGIPDLKIDFSDVPAII
    AIIYGLAGIAVNAIKNLYOYIIQGSAGVGVGAVNFIACITPILIPRALFKRINS
    ANGLAVSLGTRAMTILMSILNLYILIPATYWFILHSPALSDSALKTAVAGILPENN
    IGIIVTVFSLFIKLPWIEQORSAHLH"
    2686..2721
    /note="putative"
    /citation={3}
    complement(2726..2974)
    /gene="fer"
    /gene="fer"
    /note="ytpba; similar to B.stearotherophilus ferredoxin;
    putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /product="ferredoxin"
    /protein_id="AAC8394.1"
    /db_xref="GI:1146198"
    /translation="MAKTIYDKKTCIACGACGAAPDIYDDGSIATVTDENKGV
    VEPEVLEDMIDAEFGCPTDSIKVADEFEGDPLKE"
    3240..4298
    /gene="ytpb"
    3240..4298
    /gene="ytpb"
    /note="putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /protein_id="AAC8394.1"
    /db_xref="GI:1146199"
    /translation="MSINFDIMVLDCICKINGERSGSAVPHLFKRBSSQITQAGL
    PQTAKYFGMAKCSRSDISADLKEKHYLAPMSSDTYKKTASGAVLRAHLSRP
    NFVHGHGAIHQOAGVLMKRLSLVOVLSHKQOGSKOYIPVTRDKTLHWKRYLSRH
    TDIHMAKSLFAMLEAHLKRIENKAAQIFVSLTSHRIGYTSROLADPLKDEWYV
    IMFVASVHYFISLPECEDEALDKDLSDVHLNALTSTRTKQWQKQGFPIORIAEI
    RLKATATIEDHIVELISEHPAFMINDIVYSAEDQLQIAEFKRMRTNKKIKQIRDLGQR
    FSYFOIRLALTKOVQOYD"
    4291..5781
    /gene="recs"
    /gene="recs"
    4291..5781
    /gene="recs"
    /function="DNA repair and homologous recombination"
    /note="similar to Escherichia coli recQ gene product;
    putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /product="DNA or RNA helicase, DNA-dependent ATPase"
    /protein_id="AAC83947.1"
    /db_xref="GI:1146200"
    /translation="MTKLQQTLYQFFGFTSEKKGQDIIESTLSGKDTIAMLPTGGGK
    SICYOLPGYMDGWLIVSPLSLMEDVOQOLKARGEKRAALNSMLNQEQVLEH
    IHRKFLYLSPALOSPVLKLSYSLFVYIDAHNCISEGHDPSPYSLGQKRL
    KLGHPPVALTATATRETLQDVNMLLEHAHRLSVNPNIALREVNAAPTAEKID
    RYIOLVENOGRGIVYCPTRKMAKELAGIKRKTSSRADPFGGLSGSDRIILQOOFI
    HNOLDVICCTNAFGKGVKPDRIYVYHPLPOTAEFMQEIQRAGDGPSPYSILRA
    PGDFELQEOITQMESVTAELADVIRVLEKTERDERDRULQVLOGVGETQARMYIH
    LPMQKTSVELMKLSYRMELKLEMHVSFLQDGGCLRALLLTYPESEYEPDGN
    LPCSHCGFDLLEQKGRSKMAPLDSWSELHRIFSIQTVGEIN"
    5841..6410
    /gene="ytpb"
    5841..6410
    /gene="ytpb"
    /note="putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /protein_id="AAC83948.1"
    /codon_start=1
    /transl_table=11
    /protein_id="AAC8394.1"
    /db_xref="GI:1146201"
    /translation="MTQLLDIFAAAAAGLFFFEEDVRDYLKMDRIDMRITWGVSTAV
    IYILADMAVMKWPSPSHYDDGINKIFSKRSIPHIFLTLLIAFEEMLFQGVLOTH
    IGLWTSILIFALHFRLYSKMLLEIIVWTAISFLGLMYEMTGNLFVPMTHAFLIDAVE
    ACOIREHVRBROKHDEHVESREKKSPESL"
    6361..7083
    /gene="ytpb"
    6361..7083
    /gene="ytpb"
    /note="putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /protein_id="AAC83949.1"
    /db_xref="GI:1146202"
    /translation="MTNMSRVERRKAONLYEDONALADDDYDDGESLPTROSVKNQR
    EOKKQKQKTKPLFVLAIVFVFPVIVLTLEFKSHPDNDHDDYEDVDISQSKYE
    VYPKSEDKNDTADTKETALQKESKREDSKREQTAARKOTAVAEKEDSPNKEAT
    ANAASSOSTVQOQDQPAEPVOVFPKRVVKNHIVQKETLIRISMKYKSRGTBEKIRA
    YNHLNGNDVYTGQVIDIPLME"
    7146..7589
    /gene="ytpb"
    7146..7589
    /gene="ytpb"
    /note="putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /protein_id="AAC83950.1"
    /db_xref="GI:1146203"
    /translation="MESTLWNLQDQFTDAPFTKQMLQALVYRKQKFPENYAAKRRMRKAS
    LILGLGLCVMIMKSPEDQLIQEILSHFYLFMLATFAVCTSYFAFKRKEKSETD
    FHLRCEIIOKSDILMPQDPDKKARSEVFMHMKHYDINLYFESEK"
    7597..7619
    /note="putative"
    /citation={3}
    /gene="ytpb"
    7675..8454
    /gene="ytpb"
    7675..8454
    /gene="ytpb"
    /note="putative"
    /citation={3}
    /codon_start=1
    /transl_table=11
    /protein_id="AAC83951.1"
    /db_xref="GI:1146204"
    /translation="MKLSYKIAVGVLVAAAMTAKMYATAKGNHLKTHNPLSKMGK
    PLPTTFEISDIHKLRIIDDLLEKARSHAPHLVIGDLAEGVSPSARIEENTKRLVHF
    GYPIYVWGNNDYEVROHKLYSIFRAHGYITLIRNESVPSYNGHTIATAGVDIIRRM
    DHVEAIRKELDESQNLIVCHNPEIHQIENDDGIDVILSGHTHGQIRFGKFGPEL
```

```

Query Match 51.3%; Score 198.4; DB 1; Length 24887;
Best local Similarity 71.2%; Pred. No. 6.7e-35;
Matches 262; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

OY 20 AAAATGAGAAACAAGCTGATCTAACAAGTCACCAAGACCTAAACATGTGACCTGATAT 79
    ||||| ||| ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 21672 AAAATCTCTGCAACGCGATCCGATTAATCAACCAACGACCTGCTGTTCCCGAAAT 21613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 80 CGGATCGGTTACCAACAACTCTTAATGAATGAAGAAATATACAGCTGTACGATTTCCAT 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21612 TGGATACGGCTGCGCAAAACCTGACGAATTAATATATATGCTCTCAATTTCTGT 21553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 140 ATTATCTACTTTGTTAAACGCTAATGACACCGGTTTTTGAATTTGTATAAATTTGAGC 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21552 GTTATCAGTTTATTAACCGCTAATAACACAGGCTTTTGTGCGGTACAAAATTTTGGC 21493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 200 GACCAATTCATCGCTTGTGTCATCTTCACGACGTTAACCATATAAATAATACATC 259
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21492 CACTTCTCATCGACGCTGTACGCGCTTACAGGCGCTTACACATATAATCATCGTC 21433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 260 CGCTTATCATGCGCATTTTGCCTGCGCTCAATTTTGTGTTGAATGCGCATCAC 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	21432	CGCTTCATCATTGGCGATTTCAGCTTGCTGGCGAATCTGCCCTAAACGGCTCATCAC	21373
Oy	320	AATTCAATACCACCTGTATCAATAATATGTGAATCATGTGTAAACCAATTCACCTGAAGA	379
Db	21372	GATATCAATACCGCCGCTATCATCAATTAATAATCATTAATTCAGCCATTCAGCCGAGCT	21313
Oy	380	ATAAATAC	387
Db	21312	GTAATCC	21305

Search completed: September 12, 2003, 19:27:28
Job time : 1589.02 secs

THIS PAGE BLANK (USPTO)

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Tamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR P-PSDB: AAU33787.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27: Seq ID No 4228; 511bp; English.

XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 1305 BP; 447 A; 181 C; 277 G; 400 T; 0 other;

Query Match 100.0%; Score 1305; DB 23; Length 1305;
Best Local Similarity 100.0%; Pred. No. 5.2e-290;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACTTAACCTTATGCTATTTAGTGGAGCCCTAGTAAATCTACAAATTTT 60
DB 1 ATGACTTAACCTTATGCTATTTAGTGGAGCCCTAGTAAATCTACAAATTTT 60
QY 61 AATAGAAATGTTGAGAGCGTGTGATGTGAAGATGGCAGGTGAACAGCAT 120
DB 61 AATAGAAATGTTGAGAGCGTGTGATGTGAAGATGGCAGGTGAACAGCAT 120
QY 121 CGTATTTATTTCTCAGGTGAGTGTGAACATGATTTCAATATTTAGTACAGTGT 180
DB 121 CGTATTTATTTCTCAGGTGAGTGTGAACATGATTTCAATATTTAGTACAGTGT 180
QY 181 ATTGAATTTGTTGATCCACCATTTCCAAACAAATTAAGAGCGAGCAAAATCGCCATA 240
DB 181 ATTGAATTTGTTGATCCACCATTTCCAAACAAATTAAGAGCGAGCAAAATCGCCATA 240
QY 241 GATGAGCGGATGTTATTTATTTATTTATGCTTAAGTCGCTGAAGATTTGACAAAGCAT 300
DB 241 GATGAGCGGATGTTATTTATTTATTTATGCTTAAGTCGCTGAAGATTTGACAAAGCAT 300
QY 301 GAATGCTGCTCAAAATTTTATCAAAATCTAAAAAAACGGTCTATTTAGCGTTAAACAA 360
DB 301 GAATGCTGCTCAAAATTTTATCAAAATCTAAAAAAACGGTCTATTTAGCGTTAAACAA 360
QY 361 GTAGATTAATTTGAGAAATGCGTACAGAGCTGTATGATTTCTATTCATTAGAGATTTGGTAA 420
DB 361 GTAGATTAATTTGAGAAATGCGTACAGAGCTGTATGATTTCTATTCATTAGAGATTTGGTAA 420
QY 421 CGGTATCAATATCAGAGGTACATGTTAGTCTGTGGTGTGATGTTAGTACAGTGTGT 480
DB 421 CGGTATCAATATCAGAGGTACATGTTAGTCTGTGGTGTGATGTTAGTACAGTGTGT 480
QY 481 TCTCATTTTGGTGAAGAGAAAGATCTTATGATGAAGATCAATTCAGTATCCATT 540
DB 481 TCTCATTTTGGTGAAGAGAAAGATCTTATGATGAAGATCAATTCAGTATCCATT 540

QY 541 ATTGAGACGACCAAACTAGTAAATCAAGTTAGTAAATCTATTTAGTGAAGATCGC 600
DB 541 ATTGAGACGACCAAACTAGTAAATCAAGTTAGTAAATCTATTTAGTGAAGATCGC 600
QY 601 GTTATGCTTTCTAATGTTGACAGACAGACAGACCCCTATTCATTCAGATTTAGTTAT 660
DB 601 GTTATGCTTTCTAATGTTGACAGACAGACAGACCCCTATTCATTCAGATTTAGTTAT 660
QY 661 GATGACAAAGATTTATGTTTAAATCGATCTGCTGTGCTGCTAAAAAGAAAAAGTATAT 720
DB 661 GATGACAAAGATTTATGTTTAAATCGATCTGCTGTGCTGCTAAAAAGAAAAAGTATAT 720
QY 721 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGCATGTAACCTTCAATGTT 780
DB 721 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGCATGTAACCTTCAATGTT 780
QY 781 GTTTTATGCTTTATGATGCAAGAACAGCATTTATGAACAAATTAACCTGTTGACAGA 840
DB 781 GTTTTATGCTTTATGATGCAAGAACAGCATTTATGAACAAATTAACCTGTTGACAGA 840
QY 841 TATGACATGAACAGAGTAAAGCAGTGTGATGCTGTAATTAATGATGATGAGAA 900
DB 841 TATGACATGAACAGAGTAAAGCAGTGTGATGCTGTAATTAATGATGATGAGAA 900
QY 901 AAAGATTAATAAAGATGAAGAAATTTGAAGATGAAGTACGTAAAGATTTCAATTTTAA 960
DB 901 AAAGATTAATAAAGATGAAGAAATTTGAAGATGAAGTACGTAAAGATTTCAATTTTAA 960
QY 961 GATTATGCAAAATGCTTTGCTGCTGAAGAACGCAAGATTTAGTACATTTATTC 1020
DB 961 GATTATGCAAAATGCTTTGCTGCTGAAGAACGCAAGATTTAGTACATTTATTC 1020
QY 1021 CCTTACATTAATGAAGCAAGTGAAGAAACCATTAATAAAGCTGTCAAGTTCAACTTAAAT 1080
DB 1021 CCTTACATTAATGAAGCAAGTGAAGAAACCATTAATAAAGCTGTCAAGTTCAACTTAAAT 1080
QY 1081 GAAGTCTTACTGATGCAATTTCCATGACCCCTACACCAAGCAAGTACAGCTTTG 1140
DB 1081 GAAGTCTTACTGATGCAATTTCCATGACCCCTACACCAAGCAAGTACAGCTTTG 1140
QY 1141 AATGCTTTTATGCAACCAAGTGTGATGAAGCAACCGCATTTGTTGATTTGTTAAT 1200
DB 1141 AATGCTTTTATGCAACCAAGTGTGATGAAGCAACCGCATTTGTTGATTTGTTAAT 1200
QY 1201 GATGTAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAAAATCGCGCGCT 1260
DB 1201 GATGTAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAAAATCGCGCGCT 1260
QY 1261 TTTGCTTTGAGGTACCAATTCATTTATTAAGTCTGGAAGAGA 1305
DB 1261 TTTGCTTTGAGGTACCAATTCATTTATTAAGTCTGGAAGAGA 1305

RESULT 2
AAS54865
ID AAS54865 standard; DNA: 1311 BP.
XX
AC AAS54865;
XX
DT 13-FEB-2002 (first entry)
XX
DE *Staphylococcus aureus* DNA for cellular proliferation protein #1177.
XX
XX Antisense: ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS *Staphylococcus aureus*.
XX
PN M0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000: 2000US-191078P.
PR 23-MAY-2000: 2000US-206848P.
PR 26-MAY-2000: 2000US-207727P.
PR 23-OCT-2000: 2000US-242578P.
PR 27-NOV-2000: 2000US-253625P.
PR 22-DEC-2000: 2000US-257931P.
PR 16-FEB-2001: 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR P-PSDB: AANJ37006.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 8502; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1311 BP; 451 A; 181 C; 277 G; 402 T; 0 other;
Query Match 100.0%; Score 1305; DB 23; Length 1311;
Best Local Similarity 100.0%; Pred No. 5.2e-290;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTTAATGTAGTAATCTACAAATTTT 60
DB 1 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTTAATGTAGTAATCTACAAATTTT 60
QY 61 AATAGATAGTTGGAGAACGTTGTTGATTGTAAGATACCGCAGGTGTACACAGAT 120
DB 61 AATAGATAGTTGGAGAACGTTGTTGATTGTAAGATACCGCAGGTGTACACAGAT 120
QY 121 CGTATTTATTTCTCAGGTAGTGGTTAACACATGATTTCAATTTATTTGATACAGTG 180
DB 121 CGTATTTATTTCTCAGGTAGTGGTTAACACATGATTTCAATTTATTTGATACAGTG 180
QY 181 ATTGAATTTGTGATGACACATTCCAAACAAATTAGAGCCGACAGAAATCGCCATA 240
DB 181 ATTGAATTTGTGATGACACATTCCAAACAAATTAGAGCCGACAGAAATCGCCATA 240
QY 241 GATGAAGCGGATGTTATTTATTTATGTTAACGTCGTAAGAGATTGACAAACGAT 300
DB 241 GATGAAGCGGATGTTATTTATTTATGTTAACGTCGTAAGAGATTGACAAACGAT 300
QY 301 GAAATGTCGCTCAAAATTTTATCAAAATCAAAAAACCGGTCTATTAGCGGTTAAACAA 360
DB 301 GAAATGTCGCTCAAAATTTTATCAAAATCAAAAAACCGGTCTATTAGCGGTTAAACAA 360
QY 361 GTAGATTAATATGAAATGCGTACAGACGTGTATGATTTCTATTCTAGGATTTGGTGA 420

DB 361 GTAGATTAATATGAAATGCGTACAGACGTGTATGATTTCTATTCTAGGATTTGGTGA 420
QY 421 CCCTATCCAAATATCAGGTCACATGTTAGGTCCTGTCGACATGTTAGATGACAGTTGT 480
DB 421 CCCTATCCAAATATCAGGTCACATGTTAGGTCCTGTCGACATGTTAGATGACAGTTGT 480
QY 481 TCTCATTTTGTGTAAGAGAGAGATCCCTATGATGAAGATATTCATTTGCAATTCATT 540
DB 481 TCTCATTTTGTGTAAGAGAGAGATCCCTATGATGAAGATATTCATTTGCAATTCATT 540
QY 541 ATTGACGACCAACAGCTAGGTAATCAAGTTTATGTAATGCTATTTTAAAGTGAAGATCCG 600
DB 541 ATTGACGACCAACAGCTAGGTAATCAAGTTTATGTAATGCTATTTTAAAGTGAAGATCCG 600
QY 601 GTTATGCTTTCTATATGTTGACGAGACACAGAGACGCTATTTATACAGATATCTAT 660
DB 601 GTTATGCTTTCTATATGTTGACGAGACACAGAGACGCTATTTATACAGATATCTAT 660
QY 661 GATGACAAAGATATGTTTAAATGATACGTCGTATGCGTAAAAAAGAGAAATATAT 720
DB 661 GATGACAAAGATATGTTTAAATGATACGTCGTATGCGTAAAAAAGAGAAATATAT 720
QY 721 GAATCAACGTGAGAAATATTCAGTATTAAAGCTTTAAACGATTGAACGTTCAAAATGTT 780
DB 721 GAATCAACGTGAGAAATATTCAGTATTAAAGCTTTAAACGATTGAACGTTCAAAATGTT 780
QY 781 GTTTTGTGTTTATGATGACAGAAACAGGATTTATGACAGATTAACGTTGTCAGCA 840
DB 781 GTTTTGTGTTTATGATGACAGAAACAGGATTTATGACAGATTAACGTTGTCAGCA 840
QY 841 TATGCAATGACAAAGTAAGACAGTCGTATGTTGTAATTAATGAGTACTGTGGA 900
DB 841 TATGCAATGACAAAGTAAGACAGTCGTATGTTGTAATTAATGAGTACTGTGGA 900
QY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTCAAATTTT 960
DB 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTCAAATTTT 960
QY 961 GATTATGACAAATTTGCTTTGTCGTCGTAAGAACGCAAGATTTACTATATATTC 1020
DB 961 GATTATGACAAATTTGCTTTGTCGTCGTAAGAACGCAAGATTTACTATATATTC 1020
QY 1021 CCTTACATTTATATGACAGTGAAGAAACCATTAATAAAGTTCACACTTTAAAT 1080
DB 1021 CCTTACATTTATATGACAGTGAAGAAACCATTAATAAAGTTCACACTTTAAAT 1080
QY 1081 GAAATGCTTACTGATGCAATTTTCCATGAACCTTACACCAACAGACAAAGGTAGACGTTTG 1140
DB 1081 GAAATGCTTACTGATGCAATTTTCCATGAACCTTACACCAACAGACAAAGGTAGACGTTTG 1140
QY 1141 AATGCTTTTATGACACAGTTCGTTATGAAACCAACGCAATTTGTTGTTATTT 1200
DB 1141 AATGCTTTTATGACACAGTTCGTTATGAAACCAACGCAATTTGTTGTTATTT 1200
QY 1201 GATGTAATTAATGATTTTCTTATTAACGCTATTTAGAGAAATCAAAATCGTCCGCT 1260
DB 1201 GATGTAATTAATGATTTTCTTATTAACGCTATTTAGAGAAATCAAAATCGTCCGCT 1260
QY 1261 TTTGGTTTGAAGGTACACCAATTCATTTATATAGCTGGAAGAGA 1305
DB 1261 TTTGGTTTGAAGGTACACCAATTCATTTATATAGCTGGAAGAGA 1305
RESULT 3
AAF86461
ID AAF86461 standard; DNA: 1311 BP.
XX
AC AAF86461;
XX
DT 26-JUN-2001 (first entry)
XX
DE Staphylococcus aureus yphc coding sequence.

XX yphc; antimicrobial; cytostatic; antiulcer; microbial infection;
KW gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
KM gastritis; ds.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT CDS 1..1311
FT /product= "Staphylococcus aureus yphc protein"
XX
PN M0200123418-A1.
PD 05-APR-2001.
XX
PF 19-SEP-2000; 2000MO-US25566.
XX
PR 28-SEP-1999; 99US-0406968.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Zalcain M, Blawas S, Burnham MKR, Sylvester D, McDevitt D;
PI Methle TB;
XX
DR WPI; 2001-308138/32.
DR P-PSDB; AAB82089.
XX
PT Novel yphc polypeptides of Staphylococcus aureus useful for diagnosing
PT and treating microbial infections, especially infection by
PT Staphylococcus aureus and Helicobacter pylori
XX
PS Claim 2; Page 2-3; 41pp; English.
XX
CC The present sequence is the gene encoding yphc polypeptide of
CC Staphylococcus aureus. The yphc coding sequence and protein are useful
CC for treating and diagnosing microbial infections such as infection caused
CC by S.aureus and Helicobacter pylori. In addition, the yphc coding
CC sequence and protein are useful for treating diseases such as
CC H.pylori-induced cancers, e.g. gastrointestinal carcinoma, gastric
CC ulcers, and gastritis. The present sequence was obtained from a library
CC of clones of chromosomal DNA of S.aureus in E.coli. The sequencing data
CC from two or more clones comprising overlapping S.aureus DNAs was used to
CC construct the present contiguous DNA sequence.
XX
SQ Sequence 1311 BP; 451 A; 184 C; 278 G; 398 T; 0 other;
Query Match 99.0%; Score 1297.2; DB 22; Length 1311;
Best, Local Similarity 99.4%; Pred. No. 4; Se-287;
Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATAGCTATTTGATGAGCGCTAATGAGTAATCTACAATTTT 60
DB 1 ATGACTAAACCTATAGCTATTTGATGAGCGCTAATGAGTAATCTACAATTTT 60
QY 61 AATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAAT 120
DB 61 AATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAATGAGTAAT 120
QY 121 CGTATTTATCTTCAGGTGAGTGTGATTTGAAGATAGCCAGGTGTAACAGAT 180
DB 121 CGTATTTATCTTCAGGTGAGTGTGATTTGAAGATAGCCAGGTGTAACAGAT 180
QY 181 ATGGAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 ATGGAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 GATGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 GATGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GAAATGCTCGCTCAAAATTTTATACAAATCTAAAAAACCGGCTGATTAGCGTTAACAAA 360

DB 301 GAAATGCTCGCTCAAAATTTTATACAAATCTAAAAAACCGGCTGATTAGCGTTAACAAA 360
QY 361 GTACATTAATATGGAATGCGTACAGAGCTGATGATTTCTATTCTATTAGATTGGTGA 420
DB 361 GTACATTAATATGGAATGCGTACAGAGCTGATGATTTCTATTCTATTAGATTGGTGA 420
QY 421 CCGTATCCAAATATCAGGCTACATGTTAGTCTTGTGATGATGATGATGATGATGAT 480
DB 421 CCGTATCCGATATCAGGCTACATGTTAGTCTTGTGATGATGATGATGATGATGAT 480
QY 481 TCTCATTTTGTGTAAGAGGAAGAAATCTTATGATGATGATGATGATGATGATGAT 540
DB 481 TCTCATTTTGTGTAAGAGGAAGAAATCTTATGATGATGATGATGATGATGATGAT 540
QY 541 ATTGACAGCAAACTAGTAAATCAAGTTTACTAATGCTATTATTAGTGAAGATCGC 600
DB 541 ATTGACAGCAAACTAGTAAATCAAGTTTACTAATGCTATTATTAGTGAAGATCGC 600
QY 601 GTTATCGTTTCTAATGTTGACAGGACACGAGCCTATTGATACAGATATATTTAT 660
DB 601 GTTATCGTTTCTAATGTTGACAGGACACGAGCCTATTGATACAGATATATTTAT 660
QY 661 GATGACAAAGATTTATGTTTAAATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GATGACAAAGATTTATGTTTAAATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GAATCAACTGAGAAATATTCAGTATTAAGCTTTAAAGCGATTGAACGTTCAATGTT 780
DB 721 GAATCAACTGAGAAATATTCAGTATTAAGCTTTAAAGCGATTGAACGTTCAATGTT 780
QY 781 GTTTATGCTGTTATGATGACGAACAAGCATTTTGAACAGATTAACGTTGACGA 840
DB 781 GTTTATGCTGTTATGATGACGAACAAGCATTTTGAACAGATTAACGTTGACGA 840
QY 841 TATGACATGATGAACGAAGTAAAGCGCTGATTTCTTAATTAATGATGATGATGAT 900
DB 841 TATGACATGATGAACGAAGTAAAGCGCTGATTTCTTAATTAATGATGATGATGAT 900
QY 901 AAGATATGTTAAACGATGAAGAAATTTGAAGATGATGATGATGATGATGATGAT 960
DB 901 AAGATATGTTAAACGATGAAGAAATTTGAAGATGATGATGATGATGATGATGAT 960
QY 961 GATTATGACAAATTTGCTTTTGTGCTGTAAGACGACCAAGATTAGCTATTTTC 1020
DB 961 GATTATGACAAATTTGCTTTTGTGCTGTAAGACGACCAAGATTAGCTATTTTC 1020
QY 1021 CCTTACATTAATGAGAAAGTGAACCAATTAACCAATTAACCAATTAACCAATTAAT 1080
DB 1021 CCTTACATTAATGAGAAAGTGAACCAATTAACCAATTAACCAATTAACCAATTAAT 1080
QY 1081 GAAATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 GAAATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 AATGCTTTTATGACACCAAGTTGCTATGAAACCAAGCATTTGTTGATTTGTTAAT 1200
DB 1141 AATGCTTTTATGACACCAAGTTGCTATGAAACCAAGCATTTGTTGATTTGTTAAT 1200
QY 1201 GATGTAAGTAATTAATGATTTTCTTATTAACCGTATTATAGAAATCAAAATCCGCGCT 1260
DB 1201 GATGTAAGTAATTAATGATTTTCTTATTAACCGTATTATAGAAATCAAAATCCGCGCT 1260
QY 1261 TTTGCTTTTGAAGTACCAAAATTCATTTATAGCTGAAAGAGA 1305
DB 1261 TTTGCTTTTGAAGTACCAAAATTCATTTATAGCTGAAAGAGA 1305
RESULT 4
AAS54997
ID AAS54997 standard; DNA: 1311 BP.
XX
AC AAS54997;

XX 13-FEB-2002 (first entry)
XX DT
XX Staphylococcus aureus DNA for cellular proliferation protein #1309.
DE
XX Antisense: ds: prokaryotic cellular proliferation gene;
XX antibiotic: antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-263088P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX P-PSDB: AAU37138.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 8634; 511bp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 1311 BP; 452 A; 184 C; 278 G; 397 T; 0 other:
XX
XX Query Match 98.9%; Score 1290.6; DB 23; Length 1311;
XX Best Local Similarity 99.3%; Pred. No. 1.1e-286;
XX Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 121 CGTATTATCTTCAGGTGAATGGTTAAACATGATTTCAATATTATTGATACAGTGGT 180
Oy 181 ATTGAATTTGGTGAATGACACCAATCCAAACACAAATTAAGCGGACGGAATGGCCATA 240
Db 181 ATTGAATTTGGTGAATGACACCAATCCAAACACAAATTAAGCGGACGGAATGGCCATA 240
Oy 241 GATGAGCGGATGTTATATTATTTTATGTTAAAGCGCTGAAGAGATTAACAAAGCAT 300
Db 241 GATGAGCGGATGTTATATTATTTTATGTTAAAGCGCTGAAGAGATTAACAAAGCAT 300
Oy 301 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAAACCGGTGATTAGCGGTTAACAA 360
Db 301 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAAACCGGTGATTAGCGGTTAACAA 360
Oy 361 GTAGTATATATGGAATTCGTACAGACCTGTATGATTTCTATCATTTAGATTTGGTGA 420
Db 361 GTAGTATATATGGAATTCGTACAGACCTGTATGATTTCTATCATTTAGATTTGGTGA 420
Oy 421 CCGTATCCAAATATCAGGGTCAATGTTAGTCTTGGTGAATTTAGATGACAGTTGTT 480
Db 421 CCGTATCCGATATCAGGGTCAATGTTAGTCTTGGTGAATTTAGATGACAGTTGTT 480
Oy 481 TCTCATTTTGGTGAAGAGAGAAAGATCCTTATGATGAAGATACAAATTCGACTATCAT 540
Db 481 TCTCATTTTGGTGAAGAGAGAAAGATCCTTATGATGAAGATACAAATTCGACTATCAT 540
Oy 541 ATTGACGACCAACCAAGTATGATTAATCAAGTTTATGATTAATTTAGTGAAGATCGC 600
Db 541 ATTGACGACCAACCAAGTATGATTAATCAAGTTTATGATTAATTTAGTGAAGATCGC 600
Oy 601 GTTATCGTTCTTAATGTTGACGAGCAACGAGACGCTATTGATACAGAGTATGTTAT 660
Db 601 GTTATCGTTCTTAATGTTGACGAGCAACGAGACGCTATTGATACAGAGTATGTTAT 660
Oy 661 GATGACAAAGATTTATGTTTATATCGATACCTGCTGATGCTTAAAGAGAAAGATATAT 720
Db 661 GATGACAAAGATTTATGTTTATATCGATACCTGCTGATGCTTAAAGAGAAAGATATAT 720
Oy 721 GAATCACTGGAATATATTCATGTTATAGACTTTTAAAGGATTTGAACGTTCAATGTT 780
Db 721 GAATCACTGGAATATATTCATGTTATAGACTTTTAAAGGATTTGAACGTTCAATGTT 780
Oy 781 GTTTAGTGGTTATGATGACGAACAGCAATTTATGAACAAGATTAACGTTGACAGA 840
Db 781 GTTTAGTGGTATGATGATGACGAACAGCAATTTATGAACAAGATTAACGTTGACAGA 840
Oy 841 TATGCACATGAACAAGTAAAGACGTGATGTTGCTAAATTAATGGGATCTGTGAA 900
Db 841 TATGCACATGAACAAGTAAAGACGTGATGTTGCTAAATTAATGGGATCTGTGAA 900
Oy 901 AAAGATAGTAAAGCATGAAGAAATTTGAAGATGAGTAAAGATTTGCCAATTTTGA 960
Db 901 AAAGATAGTAAAGCATGAAGAAATTTGAAGATGAGTAAAGATTTGCCAATTTTGA 960
Oy 961 GATTATGACAAATTTGCTTTGTGCTGCTTAAGAACGCAAGATTTACGATCAATTTTC 1020
Db 961 GATTATGACAAATTTGCTTTGTGCTGCTTAAGAACGCAAGATTTACGATCAATTTTC 1020
Oy 1021 CCTTACATTAATGAAGCAAGTAAAGCAATTAAGCAAGTAAAGCAAGTAAAGCAAGTAA 1080
Db 1021 CCTTACATTAATGAAGCAAGTAAAGCAATTAAGCAAGTAAAGCAAGTAAAGCAAGTAA 1080
Oy 1081 GAAGTCGTACTGATGCAATTTTCAATGAACCTTACCAAGCAAGCAAGCAAGTAAAG 1140
Db 1081 GAAGTCGTACTGATGCAATTTTCAATGAACCTTACCAAGCAAGCAAGCAAGTAAAG 1140
Oy 1141 AATGCTTTTATGCAACCAAGTTGCTTATGAACACCGACGATTTGTTATTTGTTAAT 1200
Db 1141 AATGCTTTTATGCAACCAAGTTGCTTATGAACACCGACGATTTGTTATTTGTTAAT 1200
Oy 1201 GATGTAAGATTAATGATATTTTCTTATTAACGCTATTTAGAGATCAATCCGTGCGCT 1260
Db 1201 GATGTAAGATTAATGATATTTTCTTATTAACGCTATTTAGAGATCAATCCGTGCGCT 1260

QY 1261 TTGTGTTTGAAGTACCAATTCATATTAGCTCGAAGAAGA 1305
 |||||||
 DB 1261 TTGTGTTTGAAGTACCAATTCATATTAGCTCGAAGAAGA 1305

RESULT 5
 AA774669
 ID AAV74669 standard; DNA; 3621 BP.

XX AAV74669;
 XX
 DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #358.
 XX
 XX Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.

FH Key Location/Qualifiers
 FT misc-feature 481..540
 FT /tag= a
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"
 FT misc-feature 2281..2340
 FT /tag= b
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"

XX EP786519-A2.
 PD 30-JUL-1997.
 XX
 XX 07-JAN-1997; 97EP-0100117.
 PF
 XX 05-JAN-1996; 96US-0009861.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 PI
 XX WPI; 1997-374922/35.
 DR
 XX
 XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 PS
 XX Claim 1: Page 1241-1243; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S. aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S. aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences

CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S. aureus DNA sequences contained on the
 CC computer readable medium.
 XX

SO Sequence 3621 BP; 1279 A; 442 C; 692 G; 1085 T; 123 other;

Query Match 94.3%; Score 1230.2; DB 18; Length 3621;
 Best Local Similarity 94.6%; Pred. No. 1e-272;
 Matches 1235; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGCTATATGTAGGTAGCCCTAATGTAGTAAATCTCAATTTT 60
 |||||||
 DB 2130 ATGACTAAACCTATAGCTATATGTAGGTAGCCCTAATGTAGTAAATCTCAATTTT 2189

QY 61 AATGAGATAGTTGGAGACGCTGTTTCATTTGTTGAAGATACGCCAGTGTAAACAGAGAT 120
 |||||||
 DB 2190 AATGAGATAGTTGGAGACGCTGTTTCATTTGTTGAAGATACGCCAGTGTAAACAGAGAT 2249

QY 121 CGTATTATTCCTTCAGCTGAGTGGTTTAAACATGATTTCAATATTTATTCATCAGGTGCT 180
 |||||||
 DB 2250 CGTATTATTCCTTCAGCTGAGTGGTTTAAACATGATTTCAATATTTATTCATCAGGTGCT 2309

QY 181 ATTGAATTTGGTATGCACCATTTCCAAACACAATTTAGAGCGCAGAGAAATCGGCATA 240
 |||||||
 DB 2310 NNN 2369

QY 241 GATGAAGCGGATGTTATTTATTTATGTTTAACTGTCGCGTGAAGATTTGACACAAACGAT 300
 |||||||
 DB 2370 GATGAAGCGGATGTTATTTATTTATGTTTAACTGTCGCGTGAAGATTTGACACAAACGAT 2429

QY 301 GAAATGCTGCTCAATTTTATTCATCAATCTAAAAACCGCTGTTATTCGCTTAAACAA 360
 |||||||
 DB 2430 GAAATGCTGCTCAATTTTATTCATCAATCTAAAAACCGCTGTTATTCGCTTAAACAA 2489

QY 361 GTAGATATATATGAAATGCGTACAGACGCTATGATTTGTTTCAATTTAGATTTGGTGA 420
 |||||||
 DB 2490 GTAGATATATATGAAATGCGTACAGACGCTATGATTTGTTTCAATTTAGATTTGGTGA 2549

QY 421 CCGTATCAATATACAGGTCACATGTTTAAAGTCTTGTGATCTGTTAGATGACATTTGTT 480
 |||||||
 DB 2550 CCGTATCAATATACAGGTCACATGTTTAAAGTCTTGTGATCTGTTAGATGACATTTGTT 2609

QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATTAATTTGACATATTCATT 540
 |||||||
 DB 2610 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATTAATTTGACATATTCATT 2669

QY 541 ATTGAGACGACCAAGCTAGTAAATCAATGATTTAGTAAATGCTATTAGGTAAAGTCCG 600
 |||||||
 DB 2670 ATTGAGACGACCAAGCTAGTAAATCAATGATTTAGTAAATGCTATTAGGTAAAGTCCG 2729

QY 601 GTTATCGTTTCTAATGTTTCAGGACAGACAGAGACGCTATTGATACAGAGTATGTTAT 660
 |||||||
 DB 2730 GTTATCGTTTCTAATGTTTCAGGACAGACAGAGACGCTATTGATACAGAGTATGTTAT 2789

QY 661 GATGACAAAGATTATGTTTAAATCGATCTGCTGGTATGCGTAAAAAAGAAAGATATAT 720
 |||||||
 DB 2790 GATGACAAAGATTATGTTTAAATCGATCTGCTGGTATGCGTAAAAAAGAAAGATATAT 2849

QY 721 GAATCAACGAGAAAAATTCAGATTAATCAAGCTTTAAAGCGATTTGAAGCTTCAATTTT 780
 |||||||
 DB 2850 GAATCAACGAGAAAAATTCAGATTAATCAAGCTTTAAAGCGATTTGAAGCTTCAATTTT 2909

QY 781 GTTTATGCTTATTTATGATCAGACAAAGCATTTATTTGAACAAGATTAACGCTGTGACGA 840
 |||||||
 DB 2910 GTTTATGCTTATTTATGATCAGACAAAGCATTTATTTGAACAAGATTAACGCTGTGACGA 2969

QY 841 TATGCATGAAACAGATTAAGCAGTCTGATTTGCTAAATTAATGGATCTGTGGAA 900
 |||||||
 DB 2970 TATGCATGAAACAGATTAAGCAGTCTGATTTGCTAAATTAATGGATCTGTGGAA 3029

QY 901 AAAGATAGTAAAAAGATGAAGAAATTTGAAGATGAAGTGTAAAGAAATTTCAATTTTAA 960
 |||||||
 DB 3030 AAAGATAGTAAAAAGATGAAGAAATTTGAAGATGAAGTGTAAAGAAATTTCAATTTTAA 3089

OY 961 GATTATGCAAAATTCCTTTTGTCTGCTAAAGACGACAAAGATTACGATCATTTATTC 1020
|||||
DB 3090 GATTATGCAAAATTCCTTTTGTCTGCTAAAGACGACAAAGATTACGATCATTTATTC 3149
OY 1021 CCTTATCATTAATGAAGCAAGTGAAAACCATATAAAAGCGTTTCAAGTTCAACTTTAAAT 1080
|||||
DB 3150 CCTTATCATTAATGAAGCAAGTGAAAACCATATAAAAGCGTTTCAAGTTCAACTTTAAAT 3209
OY 1081 GAAGCGTCTACTGATGCAATTTCCATGAAACCTTACACCAACGACAAAGGTAGAGCTTTG 1140
|||||
DB 3210 GAAGTGTCTACTGATGCAATTTCCATGAAACCTTACACCAACGACAAAGGTAGAGCTTTG 3269
OY 1141 AATGCTTTTATGCAACCAAGTTGCTATAGAACGACGACATTTGTTGATTTGTTAAAT 1200
|||||
DB 3270 AATGCTTTTATGCAACCAAGTTGCTATAGAACGACGACATTTGTTGATTTGTTAAAT 3329
OY 1201 GATGTAGAAATTAATGCAATTTTCTTATTAAGCGTATTTAGAGAAATCAAAATCCGTCGCT 1260
|||||
DB 3330 GATGTAGAAATTAATGCAATTTTCTTATTAAGCGTATTTAGAGAAATCAAAATCCGTCGCT 3389
OY 1261 TTTGGTTTGAAGTACACCAATTCATATTAATGCTCGAAAGAGA 1305
|||||
DB 3390 TTTGGTTTGAAGTACACCAATTCATATTAATGCTCGAAAGAGA 3434

RESULT 6

ABN90883
ID ABN90883 standard; DNA; 1332 BP.

AC ABN90883;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:346.

KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
antibacterial; gene therapy; gene; ds.

OS Staphylococcus epidermidis.

XX US6380370-B1.

PN 30-APR-2002.

PD 13-AUG-1998; 98US-0134001.

PF 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

DR WPI; 2002-381255/41.

XX P-PSDB; ABP38338.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
polypeptide, useful for diagnosing and treating bacterial infections -
Disclosure; SEQ ID 346; 267Pp; English.

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
frame (ORF) nucleic acid sequences which encode the amino acid sequences
given in ABP35124 to ABP37960. The S. epidermidis sequences have
antibacterial activity and can be used in gene therapy. The sequences
can also be used in the diagnosis and treatment of bacterial infections,
particularly S. epidermidis infections. The sequences can be used to
screen for compounds able to interfere with the S. epidermidis life
cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from the
USPTO web site.

CC

XX
SQ Sequence 1332 BP; 465 A; 190 C; 267 G; 410 T; 0 other;

Query Match 70.6%; Score 921; DB 24; Length 1332;

Best Local Similarity 81.6%; Pred. No. 7.7e-202;

Matches 1065; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

OY 1 ATGACATAAACCTATATGATCTATTTAGTACGCTTAATGTAGTAATCTCAATTTT 60
|||||
DB 22 ATGACTAAACCATATGTATGCAATTTAGTAAACCAATGTAGAAATCTCAATTTT 81
OY 61 AATGAAATAGTTGGAAGAACGTTGTTGATTTGGAAGATACCCAGCTTACAGAGAT 120
|||||
DB 82 AATGAGTTGTGGGGAACGTTGATCTATCTATGTAAGAAATAGCCAGGCTTACTGTAT 141
OY 121 CGTATTTATTTCTTACAGTGAGTGTGTTAACACATGATTTCAATATTTAGTACAGTGT 180
|||||
DB 142 CGTATTTATTTCTTGGCGAGTGTGTTAACTCATGTAATTTAACATCATTTGATACAGTGT 201
OY 181 ATTGAATTTGGTATGACACCATTTCCAAACCAAAATTTAGAGGCGAGCAAAATCGCAT 240
|||||
DB 202 ATTGAATTTGGAGATGCTCTTTTCAAACTCAAAATCGTGACAGAGCAAAATAGCAAT 261
OY 241 GATGAAGCGGATGTTATTTTATTTTATGTTAACTGCTGAAGAGATTGACAAAGCAT 300
|||||
DB 262 GAAGAAGCAGATGTCATCATTTTATGTCATATGTAGAAAGAGACTTACACAAAGTAC 321
OY 301 GAAATGTCGCTCAAAATTTTATCAAAATCTAAAAAACCGTGATTTAGCGTTAAACAA 360
|||||
DB 322 GAAATGTCGCAAAATGCTTTTATTAAGAAACCGTTGTTTACTCTGTGAATAA 381
OY 361 GTAGATTAATATGGAATGCGTACAGCGTGTATGATTTCTATTCATTTAGTTGGTGA 420
|||||
DB 382 GTTGATTAATCTTGAATATGCTATATATCTATGATTTCTATTTCTTTAGCGTTGAGAT 441
OY 421 CCGTATCCAAATATACAGGCTCACATGTTAGTCTTGTGACTGTATGATGACAGTTGTT 480
|||||
DB 442 CCATATCCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 501
OY 481 TCTCATTTTGGTGAAGAGAGAGAAAGATCTTTATGATGAAGATACATTTGATATTCAT 540
|||||
DB 502 GAAATCTTTAATTAAGAAATCAAGAAATCTTATGAGAAATGATGATGCTTTTCCATC 561
OY 541 ATTGACGACCAAGAGTATGTAATCAAGTTAGTAATGATTTAGTGAAGTATGCG 600
|||||
DB 562 ATCGTAGACCTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 621
OY 601 GTTATCGTTTCTAATGTTGTCAGGACCAAGAGAGCGTATTTAGTATACAGATATAGTTAT 660
|||||
DB 622 GTTATTTGTCTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 681
OY 661 GATGACAAAGATTAATGTTTAAATGATATGCTGTTATGCGTTAAAAAGAAAGTATAT 720
|||||
DB 682 GATGACAAAGATTAATGTTTAAATGATATGCTGTTATGCGTTAAAAAGAAAGTATAT 741
OY 721 GAATCACTGGAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 780
|||||
DB 742 GAATGACTGGAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 801
OY 781 GTTATGAGTTTATGATGACAGAACAGCATTAATTAAGAAAGATTAATGATTAATGAT 840
|||||
DB 802 GTATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
OY 841 TATGACATGAACAAGTGAAGCACTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
|||||
DB 862 TATGACATGAACAAGTGAAGCACTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 921
OY 901 AAGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 960
|||||
DB 922 AAGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 981
OY 961 GATTATGCAAAATTCCTTTTGTCTGCTAAAGACGACAAAGATTACGATCATTTATTC 1020
|||||

CC	Proteins (AB847297-AB850149) expressed from the present sequence are
CC	useful for raising specific antibodies, identification of L.
CC	monocytogenes and related organisms, and for biosynthesis and
CC	biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC	proteins encoded by it are also useful for selecting compounds that
CC	regulate gene expression and cell replication and modulate L.
CC	monocytogenes-related diseases. In addition, this sequence and proteins
CC	encoded by it are useful in pharmaceutical and vaccines compositions for
CC	the treatment or prevention of infections by L. monocytogenes and related
CC	organisms.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/Published_pcl/sequences .
XX	
SO	Sequence 2944528 BP: 914202 A: 563301 C: 555061 G: 911964 T: 0 other;
Query Match	52.4% Score 683.2: DB 24: Length 2944528:
Best Local Similarity	70.2% Pred. No. 2.3e-146:
Matches	916: Conservative 0: Mismatches 388: Indels 0: Gaps 0:
QY	1 ATGACTAAACCTATAGTAGCTATGTAGTAGAGCCTAATGTAGGTAATCTACAATTTT 60
DB	2012438 ATGGCAAAACCAAGTGTAGCATTTGTCGGACGTCGCAAAACGTTGGCAATTCAGCATTTT 2012379
QY	61 AATAGAAATAGTTGGGAACGCTTTCCATTGTTGAAGTACGCCACAGCTGAACACAGAGT 120
DB	2012378 AACACAAACGTTGGGAACGCTTTCCATTGTTGAAGTATGTTCCCGGTGTACACAGTGC 2012319
QY	121 CGTATTTATTTCTTCAGCGTAGAGGGTTTAACACATGTTTCANATTTATGTATACAGGTGT 180
DB	2012318 CGCATATTTAATTCACGCCGAATGCGTTGGAAAAGCAATTTAACTATTATGTATACAGGTGT 2012255
QY	181 ATTGAAATTTGGTGTATGCACATTTCCAAACACAAATTTAGAGCGACAGCAAGTAATCGCCATA 240

[illegible]

DB 2011658 GTCTTGGGTATACGACGACAAAGGATATTCGTGACGACGATTAAGCAATTCCTGGA 2011599
QY 841 TATGACATGAACAGTAAAGCAGTCGATGTCGTAAATAATGGATCTGTGAA 900
DB 2011598 TATGGCATGATGCGGAGCGTCATATATATGTTGGAACAATGGATCATTAAAC 2011539
QY 901 AAAGTACTAAACCATGACGAAATTTGAAGATGAAGTAAAGTAAAGATTCGAATTTTA 960
DB 2011538 AAAGATGAAAAACAATTAATGATAGACAGAAAGATATTCGTGACCAATTCGAATTCCTTA 2011479
QY 961 GATTATGACAAATGCTTTTGTCTGTAAAGACGACAAAGATTAAGTACATTAATTC 1020
DB 2011478 AGCTATGACCAATGTTTGTCTATGCTAAACAAACAACGCTTAAATATCTATTC 2011419
QY 1021 CCTTACATTAATGAAGCAAGTGAACCATTAACCAAGTGTCAAGTTCACTTAAAT 1080
DB 2011418 CCACATTAATTAACCAAGTAAAGGATTAACCATTCATACGTCATCAAGTATGCTGAAT 2011359
QY 1081 GAATGCTTACTGATGCAATTTCCATGAACCTTACACCAACGACAAAGTAAAGTGTG 1140
DB 2011358 GATGTTATTAATGATGCTGTTGCAATGAACCATCCTATGATTAAGTAAAGTAAAGTCTG 2011299
QY 1141 AATGCTTTTATGACACAAAGTGTCTATAGAACGACGACATTTGTGATTTTAAAT 1200
DB 2011298 AAGATTTTCTATACACACAAAGTGGCGGTAAACCAACGCTTGTGATTTTAAAT 2011239
QY 1201 GATGATGAATTAATGATTTTCTTAAAGCCTTAAAGAAATCAATCCGTCGCT 1260
DB 2011238 GATCAGAACTCATGCTATTTTCTGATAGAGCTTTCTTAAGAAACCGGATTTGAAACGA 2011179
QY 1261 TTTGCTTTGAAGTACACCAATTTATATAGTCCAAAGAG 1304
DB 2011178 TTTCCATTGAGGATCGCCCAATTCGAGTAATTTGCTCGTAAGCG 2011135

RESULT 9
AB067194
ID AB067194 standard; DNA; 319630 BP.
XX
AC AB067194;
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua contig DNA sequence #7.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KM infection; ds.
XX
OS Listeria innocua.
XX
FN WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INST PASTEUR.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
DR WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators -
XX
PS Claim 5; SEQ ID 7; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences

CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from Wipo at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 6 other:
Query Match 51.9%; Score 676.8; DB 24; Length 319630;
Best Local Similarity 69.9%; Pred. No. 3,8e-145;
Matches 912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATATGCTATTTGTAGTACGCTTAATGTATGAATCTACAAATTTT 60
DB 50618 ATGGCAAAACAGTTGTAGCGATTTGCGAGTCCAAACGTTGGCAATTCACATATTTT 50677
QY 61 AATAGATAGTTGGAGAACGTTTGTGATTTGAAGATPACGCCAGGTGAACAGATAT 120
DB 50678 AACGAAATCGTTGTGAACGTTTTCATAGTGAAGATGTTCCCGGTGTACACGTGAC 50737
QY 121 CGTATTTATTCCTTCAGGTAGTGTGTTAACACATGATTTCAATATTTATGATACAGTGT 180
DB 50728 CCGATATTAATATTCAGCGGAATGCGTGTGAAAGAAATTAACATTAATGATACAGTGT 50797
QY 181 ATTGAATTTGTTGATGACCATTCGAAACACAAATTAACGCGCAAGCAATGCGCTTA 240
DB 50798 ATTGATCTTTCCGAGCAACATTTCTTAAGCAAAATTCGCGCAACAGCGAAATGCAATT 50857
QY 241 GATGAAGCGGATGATTAATTTTATTTTATGTTAACGTCGTGAAGATGACACAAAGCAT 300
DB 50858 GATGAAGCAGACGATTAATTTTATTTTATTCACATGCTGTGAAGGCTTACCGATGACAG 50917
QY 301 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAACCGGTCGATTAAGCGGTTAAACAA 360
DB 50918 GAACAAATGACAAAAATTTCTTACCGGTCTAATAAACCAATTTGTTTGGCATTAATAA 50977
QY 361 GTAGATATATGGAATATGCTACAGACGTCATATTTCTATTCATTAAGATTTGTGTAA 420
DB 50978 GTAGATTAACCCAGAAATGCGTACATGATTAATGATTTATTCCTTTGCTGTTGGTGA 51037
QY 421 CCGTATCCAAATATGAGGTCACATGTTAGTGTGCTGATGCTTTGATGACGTTGTT 480
DB 51038 CCGTATCCAAATTTCTGCTTCATGATGACTAGGCTTGTGATATGCTGATGCTGCTTGT 51097
QY 481 TCTCATTTTGTGAAGAGAGAAAGATCTTATGATGAAGATPACAAATTCGATTCATT 540
DB 51098 GCTCATTTTCCAAAAAGAGAGAGAAAGATTAACCAAGATGAACAGTAAATTTAGTTTG 51157
QY 541 ATTGACGACCAACAGTATGTAATCAAGTTTATGAATATGCTATTTTGTGAAGATGCG 600
DB 51158 ATTGTCGCGCAAAATGTTGTAATCAATCAATTTCTAAACGCACTTTTGGGAAAGCCG 51217
QY 601 GTTATCGTTTCTAAATGTTGAGGACGACAGACAGACGCTATTTGATACAGATATGAT 660
DB 51218 GTCATGTTTCTGACATTTGCGGGAACATCTGCTGATGCAATTTGATACAACTTATACGTT 51277
QY 661 GATGACAAAGATTAATGTTTAAATGATACGCTGCTGATGCTGTAAGAAAAAGAAAGATAT 720
DB 51278 GATGCTCAGATTAATGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 51337
QY 721 GAATCACTGGAATATTTCAATTAATGAACCTTTAAAGGATTTGAACGTTCAAAATGTT 780
DB 51338 GAAAGCACAGAGAAATATAGTGTGTTTACGTCCAATGAGAGCAATTTGAACGCTCGATGTT 51397
QY 781 GTTTTATGCTTATTTGATGACAAACGCAATTTATGAAACAAAGATTAACGTTGACGGA 840
DB 51398 GTCTTGTGTTATCAACGACAGAAAGATTTGCTGACGAAATTAAGCGGATTTGCTGGA 51457

	CC	contaminants in foods, or mutational analysis) and for analysis of
	CC	gene expression. Proteins encoded by the nucleic acid sequences can be
	CC	used to screen for compounds that modulate gene expression, replication
	CC	and pathogenicity of Listeria (potential therapeutic agents), also for
	CC	treating infections by Listeria, and are useful as immunogens in
	CC	anti-Listeria vaccines.
	CC	Note: The sequence data for this patent did not form part
	CC	of the printed specification, but was obtained in electronic format
	CC	directly from WIPO at ftp.wipo.int/pub/published_pct.sequences.
	XX	
S0	Sequence	3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other:
	Query Match	51.9%; Score 676.8; DB 24; Length 3011208;
	Best Local Similarity	69.9%; Pred. No. 6.7e-145;
	Matches	912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;
QY	1	ATGACTAACCTAGTACGTATTGTAGCGCTAATGTAGTAATCTACAAATTTT 60
DB	2066810	ATGGCAAAACCAGTTGTAGCGATTGTGGAGCGTCAAACGTTGGCAAAATCAGCATTTTTT 2066751
QY	61	AATGAAGTAGTGTGAGACAAGCTGTTTCGATTGTTGAAGTACGCCAGGTGTAAACAGAGAT 120
DB	2066750	AACGAGATCGTGTTGTAACCGTGTTCATGTGGAAGATGTTCCCGGTGCACACGTGCAC 2066699
QY	121	CGTATTTATTCCTCAGGTGAGTGTGTAAACATGATTTCAATATTATTATGTATACAGTGT 180
DB	2066690	CGCATATATATAATTCACGGGAATGGCTTGGAAAAGAATTTAACATATTATTGTATACAGTGT 2066631
QY	181	ATTGAATTTGCTGATGCACCATTTCCAACACAATTTAGCGCGCACGAAGATGCCATA 240
DB	2066630	ATTGATCTTTCGCCGAGAACCTTCCTTGAGCAAAATTCGCCGCAACGGGAATTCGCAATT 2066571
QY	241	GATGACGCGGATGTATTATTATTATTATTTAGCTTACGCGGTGAAGGATTGACACAAGGAT 300
DB	2066570	GATBAGCAGACGTAATATTATTATTATTATTCACATGCTGCTGAAGGGTTACCGATGCACAC 2066511
QY	301	GAATTGGTCCGCTCAATTTTATTACAAATCTAAAAACCGGCTGATTAGCGGTTAACAAA 360
DB	2066510	GAACAGTAGCAAAAATTTCTTACCGGCTAATAAACCAATGTTTTAGCGATTAAATAAA 2066451
QY	361	GTACATATATATGGAATTCGCTACAGACGCTGATGATTTTCTATTCATTAGATTTGGTGA 420
DB	2066450	GTACATATAACCCAGAAATGCGGATACAGATTATGACTTTTATCTCTGGGCTTTGGTGAG 2066391
QY	421	CCGATATCCAAATACAGGGTGCATCGATGTTAGCTTGGCTGTGACTGTATGACGATGGTT 480
DB	2066390	CCGATATCCAAATTTCTGTTCTCATGAGCATAGGGCTTGGTGATGCTTGATGCTCTTCGT 2066331
QY	481	TCTATTATTGGTGAAGAGAAGATCCCTATGATGAAGATACAAATTCGACTTCATT 540
DB	2066330	GCCTATTTTCCAAAAGAAGAGAGGAAGATACCACATGAAACAGTGAATTTAGTTTG 2066271
QY	541	ATTGAGACACCAACGTAAGTAAATCAAGTTTACTAAATGCTATTTTAAAGTAAATCGC 600
DB	2066270	ATTGCTGGCCCAATTTGTTGTTCTCATGAGCATAGGGCTTGGTGATGCTTGATGCTCTTCGT 2066211
QY	601	GTTTATCGTTTCTAATGTTGACAGGACAACGAGACGCTATTGATACAGATATTAGTTAT 660
DB	2066210	GTCAATGTTTCTGACATTTGGCGGAAACACACGCTGATSCAATGTGATACCACTTATACGTTT 2066151
QY	661	GATGACACAAGATTATGTTTAAATGATACGCTGTTAGCTTAATAAAAAGAAAAGTATAT 720
DB	2066150	GATGCTAGAGATTATGTCATGATGATTGATACACCGCAATGAGAAAACCTGGGAAAAGCTGAT 2066091
QY	721	GAATCAACGTGAGAAATATTCAGTATTAAGACCTTTAAAGGATTTGAACGTTCAATGTT 780
DB	2066090	GAAAGCACAGAGAAATATATAGTGTTTTACGTGCAATGAGAGCAATTGAACGCTCCGATGTT 2066031
QY	781	GTTTTAGTGGTATTATGATGCAGAACAGGCAATTATTGAACAAGTAAACGTTTGCACGA 840
DB	2066030	GTTTCTTTTGTGTTATCAACGCAAGAGAGTATTCGTAGACAAAGATTAACGGAATTGGTGA 2065971
QY	841	TATGCACATGTAACAGGTGAAGCGAGTGTGTATCTGTAATTAATTTAGGATCTACTGTGGAA 900

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:59 ; Search time 3513.16 Seconds
(without alignments)
9028.139 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305
Sequence: 1 atgaactaacctagtagc.....atatatagctcgaagaaga 1305

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estbba:*
2: em_estbhum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	486.6	37.3	2662	BH770770	BH770770 LMGTAG51
2	143.6	11.0	552	A1390536	A1390536 mu83b02.y
3	128.4	9.8	721	B0573090	B0573090 B0573090
4	124	9.5	1233	AV110834	AV110834 Zee mays

5	118.6	9.1	583	14	CA782602	CA782602 sac30n09.
6	108.4	8.3	431	9	AA215138	AA215138 mu83b02.r
7	99.8	7.6	579	13	B0514303	B0514303 EST621718
8	99.2	7.6	532	13	B0281519	B0281519 WHE2022.D
9	89.4	6.9	393	9	A168410	A168410 mu83b02.x
10	88.4	6.8	745	14	CD377507	CD377507 PTM02676
11	84.6	6.5	171	28	BH234869	BH234869 PE_04.x
12	83.2	6.4	1101	29	CNS00396	AL0630321 Drosoph11
13	81.4	6.2	167	28	BH234795	BH234795 MEA_B03.
14	79.6	6.1	653	13	B085888	B085888 OCB28A05.
15	78.8	6.0	501	14	N37967	N37967 10194 Lambd
16	78.8	6.0	613	10	BG299568	BG299568 HVSMEa002
17	77.4	5.9	554	14	CA938405	CA938405 sav32b06.
18	77	5.9	631	28	BH861059	BH861059 M3B1T7 we
19	76.8	5.9	727	14	CD407144	CD407144 Gm_CK3216
20	74.2	5.7	908	28	A2548467	A2548467 ENTEK30TR
21	73.2	5.6	1359	28	BH770563	BH770563 LMGTAG32
22	73	5.6	769	13	B0802107	B0802107 WHE2822_B
23	72.6	5.6	1010	14	CD102897	CD102897 AGENCOURT
24	71.6	5.5	410	14	CD474342	CD474342 nad03-2ms
25	71.6	5.5	841	14	CB996761	CB996761 AGENCOURT
26	71.6	5.5	905	28	A2550256	A2550256 ENTEV58TR
27	71.4	5.5	487	12	B1272296	B1272296 NF022D05F
28	71	5.4	538	13	BQ490035	BQ490035 60-E9427-
29	70.6	5.4	552	14	CA742475	CA742475 wrl1s-PK0
30	70.6	5.4	623	12	BJ463635	BJ463635 BJA63635
31	69.8	5.3	843	28	A2551618	A2551618 ENTDV54TR
32	69.8	5.3	931	28	BH160272	BH160272 ENTQV49TR
33	69.2	5.3	412	12	BH402985	BH402985 SLA010D07
34	69	5.3	772	12	B1920601	B1920601 EST540536
35	68.8	5.2	849	28	A2546009	A2546009 ENTFW52TF
36	68.4	5.2	912	28	A2551092	A2551092 ENTFJ22TF
37	67.2	5.1	745	28	AY079759	AY079759 AY079759
38	66	5.1	911	28	BH164417	BH164417 ENTS163TR
39	65.8	5.0	532	10	AM979775	AM979775 EST341400
40	65.6	5.0	541	13	BQ627620	BQ627620 sap33h03.
41	65.6	5.0	877	28	A2531291	A2531291 ENTBO34TR
42	65.2	4.9	571	13	BUS78174	BUS78174 sar48q07.
43	64.2	4.9	721	10	BG646649	BG646649 EST508268
44	64	4.9	350	14	CA735279	CA735279 wpl1s-PK0
45	63.6	4.9	356	14	H37009	H37009 15138 Lambd

ALIGNMENTS

RESULT 1
BH770770
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BH770770 2662 bp DNA linear GSS 01-MAY-2002
LMGTAG515 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, genomic survey sequence.
BH770770.1 GI:20373727
GSS.
Lactococcus lactis subsp. cremoris
Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
1 (bases 1 to 2662)
Bojoltin,A., Ehrlich,S.D. and Sorokin,A.
Studies of genomes of dairy bacteria Lactococcus lactis
Sci. Aliments, (2002) In press
Contact: Sorokin A
Genetique Microbienn
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is yphL (98%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 2634.

```

FEATURES
  source
    Location/Qualifiers
      1. 2662
      /organism="Lactococcus lactis subsp. cremoris"
      /mol_type="genomic DNA"
      /strain="MG1363"
      /db_xref="taxon:1359"
      /clone_lib="MG1363 Random Sequence Tag Library"
      /note="Vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L. lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."
BASE COUNT
  871 a 506 c 513 g 772 t
ORIGIN
Query Match 37.3%; Score 486.6; DB 28; Length 2662;
Best Local Similarity 65.6%; Pred. No. 5.2e-89;
Matches 743; Conservative 0; Mismatches 384; Indels 6; Gaps 2;

QY 1 ATGACATCAACCTATATAGCTATGTAGAGCGCTTAATGTAGTAAATCTACAATTTT 60
Db 1533 ATGAGCTTACCTACAGTACGCTGCTGGCGCTCAAAATGTTGAAAATCAACGATATTC 1592
QY 61 AATGAATAGTTGGAGAGCGTGTGATTTGTAAGATACGCGAGGTGAACAGAGAT 120
Db 1593 AACCGTATTCAGAGAGAACGCAATCTCAATTTGTGAAGACATTTCCAGGTGAC 1652
QY 121 CGTATTTATTTCTTCAGGTGAGTGTGAACATGATTTCAATATTTATGATACAGTGT 180
Db 1653 CGTATCTATGCCACAGAGAAATGTTAACCCGTAATTCACATCATCGACCGGTGG 1712
QY 181 ATTGAATTTGGTATGCACCATCCAAACACAAATTAAGAGCGAGCAGAAATCCGCATA 240
Db 1713 ATTTGAATTTGATTAACCTTTTCATGATGAATTCGGCACAAATCGCATG 1772
QY 241 GATGAAGCGAGTGTATTTATTTTATGTTAACGTCGTGAAGATTTGACACAAACGAT 300
Db 1773 ACCGAAGCAGATTTATTTATCGCTGCTGTGATGAGAAACAGAGATTTACTGCGCAT 1832
QY 301 GAATAGTGCCTCAAAATTTTATCAAAATCAAAACCGGTCGTATTTAGCGGTTAACAA 360
Db 1833 GAAGCAGTTGCTAATTTCTTTATCTGTACAGATAAACCTGTCTTATCGTCAATAAA 1892
QY 361 GTGATTAATTAATGAATGCGTACAGACGTGTATTTCTATTCATTAATTTGGTAA 420
Db 1893 GTGCAACACCGGAGAGCTGAATGAATTTTGTACTTCTATTCATTTGGACTGGGAT 1952
QY 421 CCGTATCAATATCAGAGTGCACATGTTAGTCTTGTGACTTGTAGTACAGTGT 480
Db 1953 CCATATCTGTATCAGCTGTTCACGGGATTTGGACAGAGGATGTACTGTATTTGTT 2012
QY 481 TCTCATTTTGGTCAAGAGAGAGATCCTTATGATGAAGATCAATTCGACTATCCATT 540
Db 2013 CAAATCTTCTCAATCAATCAAGA---AGAAAATGCAAGATGTCAATCAAGTTCAGCTTG 2069
QY 541 ATTGAGAGCAACAGAGTAAATCAAGTGTAGTAAATGCTATTTTGAAGAGATCCG 600
Db 2070 ATTTGCTGTCACAAAGCTTGTGAATCTCTTATCAATGCAATCTTGTGTGAAGACGA 2129
QY 601 GTTATCGTTTCTATGTTCAGGAGACACAGAGCGCTATTTGATACAGAGTA---TAGT 657
Db 2130 GTGATTTGCAAGCCCAATCGCGGAGACACATCGTAGGCCATTATATACACCTTCGTTGAT 2189
QY 658 TATGATGAGACAGATTTATTTTAAATGATGCTGTGTGATTCGCTAATAAAGAGAAAGTA 717
Db 2190 TCTGAAGACCAAGATTTGTCATGATGATACCGGTGATTCGCTAATCAGAGAAATTT 2249
QY 718 TATGATCAACATGAGAAATTTTCAATTAATGAAGCTTTAATAAGATGATCAAT 777
Db 2250 TACGAACACGTAATAAATTTTCAATGATGCGGTGACGTCATGACCGTTCATGAT 2309
QY 778 GTTGTATTAGTGTATTTGATGAGAAACAGCATTTATTTGAACAGATTAACGTTGTGA 837
Db 2310 ATCGTTTGTATGCTCATTAATGCTGAAGAAGAAATTCGCAATATGATATGCTATTCGCT 2369

```

```

QY 838 GGATATGACATGACAGACAGTAAAGCATGCTGATTTGCTGTAATTAATGGCTACTGTG 897
Db 2370 GGATTTTCTCATGACAGCATTAAGCATTTCTGATTTGCTCAATTAATGGGTACTCTT 2429
QY 898 GAAAGATATGTAACGATGAAGAAATTTGAAGATGAAGTACTGAAGATTCACATTT 957
Db 2430 GAAAGACATATATCAATGAAGAAATTTGAATTTGAATTTGTCACAAATTTCAATTC 2489
QY 958 TTGATTTATGACAAATTTGCTTTTGTGCTGTCTGAAGAGCAGCAAGATTTACTACATTA 1017
Db 2490 CTGATTTATGACCAATTCGTTTACGTGTGTGCTGTCAAAACAGCCACAGATTTGAATC 2549
QY 1018 TTCCCTTACATTAATGACAGCAAGTGAACCATTAATAAACGTTTCAAGTTTAACTTTA 1077
Db 2550 CCAGATATGATTTAATAAATTCACATTCGTCACAAACCTGATTTTCATTTCTGTTTG 2609
QY 1078 AATGAAGTGTGTTCTGATGCAATTTCCATGCAACCTTACACCAACAGCAAG 1130
Db 2610 AATGATGTGATTAATGATGCGTGTGCGATTAATCAACTCCGACAGATTAAGG 2662

```

```

RESULT 2
AI390536
LOCUS
DEFINITION
  AI390536 552 bp mRNA linear EST 15-MAR-2000
  mu83b02.y1 Stragene mouse melanoma (#937312) Mus musculus CDNA
  clone IMAGE:652107 5' similar to SW:YPHC.BACSU P50743 HYPOTHETICAL
  48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. ; mRNA
  sequence.
ACCESSION
  AI390536
  AI390536.1 GI:4216543
VERSION
  EST.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  Mus musculus
  ORGANISM
    Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1 (bases 1 to 552)
REFERENCE
  AUTHORS
    Maria M., Hillier L., Kucaba T., Martin J., Beck C., Wylie T.,
    Underwood K., Stepien M., Theising B., Allen M., Bowers Y., Person
    B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk R., Riller
    E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R.,
    Waterston R. and Wilson R.
    The WashU-NCI Mouse EST Project 1999
  TITLE
    Unpublished
  JOURNAL
    Contact: Maria M/WashU-NCI Mouse EST Project 1999
  COMMENT
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: mouseest@wustl.edu
    This clone is available royalty-free through LNL; contact the
    IMAGE Consortium (info@image.lnl.gov) for further information.
    This read is a RESEQUENCE of a previously sequenced mouse clone
    This read has been verified (found to hit its original self in the
    correct orientation)
    possible reversed clone: similarity on wrong strand
    MGI:397955
    Seq primer: -40RP from Gibco
    High quality sequence stop: 420
    POLYA-No.

```

```

FEATURES
  source
    Location/Qualifiers

```

```

  1. 552
  /organism="Mus musculus"
  /mol_type="mRNA"
  /db_xref="taxon:10090"
  /clone="IMAGE:652107"
  /tissue_type="melanoma"
  /dev_stage="M2 cells"
  /lab_host="SOLR (kanamycin resistant)"
  /clone_lib="Stragene mouse melanoma (#937312)"
  /note="Organ: skin; Vector: pBluescript SK-; Site 1: EcoRI
; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. From M2 cells, a highly metastatic derivative of the
K-1735 (mouse) melanoma. Average insert size: 1.0 kb.

```


FEATURES	SOURCE	BASE COUNT	ORIGIN
Location/Qualifiers	1. 1233	335 a	268 c 294 g 278 t 58 others
/organism="Zea mays"			
/mol_type="mrna"			
/db_xref="MaRaDB:633068"			
/db_xref="taxon:4577"			
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"			
<p>/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"</p>			

Query Match	9.5%;	Score 124;	DB 11;	Length 1233;
Best Local Similarity	51.1%;	Pred. No. 5.2e-15;		
Matches	362;	Conservative	0;	Mismatches 328; Indels 18; Gaps 4;

QY	616	GTTCAGGAGCAACGAGAGACGCTATTGATACAGATATAGT---	TATGATGACAAGAT	672
Db	21	GTATAGTGGGACACACGATGATGCATATGATACGATTTACTACGACGATGGGAGAG	80	
QY	673	TATCTTTTAATCGTACTCGTGTATCCGTAAAAAGGAAAGTATAT-----	GAA	723
Db	81	TACAACTCATGTTACTGCTGGGATCCGGCGCAGCAGCATATTTCTGTCGGACG	140	
QY	724	TCAACTAGAAATATTCAGTATTAAAGACCTTAAACGATTCGAACCTTCAATGTTGT	783	
Db	141	ACAACGCAATCATTGGTTAAACGTCGATTCGACGAATTCGGCCCTCGATGTGTT	200	
QY	784	TTAGTGGTATTATGTCGCAAGAACGCAATTTATGACACAGATTAACGCTGTTGCAGATAT	843	
Db	201	GCCCTTTATTATGAGCAATGGCCCTGTCTCACAGCAGCATTTAAATATGNNNNNNNN	260	
QY	844	GCACATCAACAAAGTAAAGCAGTGTGTAATTCGCTAAATTAATGATGCTGCGAATAA	903	
Db	261	NNNNNNNNNNNNNNNNNNNNNTTGTCTATCTTGTGTAACCAATGCGATACATCCCAAC	320	
QY	904	GATAGTAAAAAGATGAAGA--AAATTTGAAGATGAAGTACGTAAAGAATTCCAATTTTGA	960	
Db	321	AAGAACCACGAGATACACACATTTATGAACAAAGATTAAGAGAAAGCCTTCGATACTT	380	
QY	961	GATATGACACAATAGCTTTTGTGTCTGCTAAAGAACGCACAAATTAACGATATTCGATATTC	1020	
Db	381	GATTGGGCACCTATTATGCTACTGTTCTGCGCAAAATGCGACACAGTATGTTGAAGACATATT	440	

RESULT	5
CA782602	
LOCUS	583 bp mRNA linear EST_04-DEC-2002
DEFINITION	sa630h09.y1 Gm-cl056 glycine max cDNA clone SOTIBEAN CLONE ID: Gm-cl056-3065 5' similar to TR:P72548 P72548 HYPOTHETICAL PROTEIN ;', mRNA sequence.
ACCESSION	CA782602
VERSION	CA782602.1 GI:26044687
KEYWORDS	EST.
SOURCE	glycine max (soybean)

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Glycine max					
Euarvota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: rosids					
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;					
Glycine.					
1 (bases 1 to 583)					
Shoemaker, R., Kelm, P., Vocklin, L., Erpelding, J., Corvelli, V., Khana					
, A., Bolla, B., Marri, M., Hiller, L., Kucaba, T., Martin, J., Beck, C.,					
Wylle, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers					
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk					
, R., Ritten, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann					
, R., Waterston, R. and Wilson, R.					
Public Soybean EST Project					
Unpublished					
Contact: Shoemaker R/Public Soybean EST project					
Public Soybean EST Project					
Washington University School of Medicine					
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA					
Tel: 314 286 1800					
Fax: 314 286 1810					
Email: est@wustl.edu					
This clone is available through: ResGen, Invitrogen Corp. 2130					
South Memorial Parkway Huntsville, AL 35801 For further information					
call: (800)-533-4363 or contact: ccu@resgen.com web site:					
www.resgen.com					
Seq primer: ~40RP from Gibco					
High quality sequence stop: 421.					
Location/Qualifiers					
1..583					

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl056-3065"
/tissue_type="Whole seedling, 4 day old"
/lab_host="DH10B"
/clone_lib="Gm-cl056"
/notes="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from a 4 day old seedling of PI68916. The seedlings were germinated in a growth chamber using germination paper. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction

```


LOCUS	BO514303	579 bp	mRNA	linear	EST 07-MAR-2003
DEFINITION	EST621718 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STM112 5' end, mRNA sequence.				
ACCESSION	BO514303				
VERSION	BO514303.1				
KEYWORDS	GI:21373172				
REFERENCE	EST.				
AUTHORS	Solanum tuberosum (potato)				
ORGANISM	Solanum tuberosum				
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Asperidites; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.				
JOURNAL	1 (bases 1 to 579)				
COMMENT	Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A. Generation of a set of potato cDNA clones for microarray analyses Unpublished Other-ESTs: EST621719 Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/Orders/ Seq primer: T3.				
FEATURES	Location/Qualifiers				
source	1..579				
	/organism="Solanum tuberosum"				
	/mol_type="mRNA"				
	/cultivar="Kennedec or Binje"				
	/db_xref="taxon:4113"				
	/clone="STM112"				
	/tissue_type="mixed tissues"				
	/lab_host="SOLR"				
	/clone_idb="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"				
	/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2: XhoI; supplier: Combination of untreated and phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."				
BASE COUNT	179 a	95 c	153 g	152 t	
ORIGIN					
Query Match	7.6%	Score 99.8:	DB 13:	Length 579:	
Best Local Similarity	55.7%	Pred. No. 4.2e-10:			
Matches 238:	Conservative	0:	Mismatches 177:	Indels 12:	Gaps 2:
QY	531	ACTATCCATTTATTTGGACGACCAACGATGAATTAATCAAGTTTAAATGCTTATTTAGG	590		
Db	21	AATCGCAATTTGTGGCGGTCGCAATGTTCGTAAGAAGTAGATTTTGAATGCTTATGTCGG	80		
QY	591	TGAAGATCGCTTATTCGTTCTTAATGTTTGCAGGACAAACGAGAGCGCTATGTATACAGA	650		
Db	81	AGAGGACAGAACAAATTTGTAGTCCCGTTAGTGGAACTACTCGTATCTATAGATACAGA	140		
QY	651	GT---ATAGTTATGATGGACAAGATTATGTTTAAATGATCTCTGCTATGCTTAATAA	707		
Db	141	ATTCACTGGACAAATGCGGACGAAGTTTGGCTTATGATTAATCTCCGGAATAAATAAAAA	200		
QY	708	AGGAAAGTATATGATC-----AATCGAAGAAATATTCAGTATTAAGAGCTTTAA	758		
Db	201	GATGCACTGACTCTATCAGGTAGCATCCAGAGAGCGTTTATCACTAATCAAGCATTTCC	260		
QY	759	AGCGATTGAACGTTCAATGTTGTTTATGTTTATGTTTATGATGACAGAACAGCATTTATGA	818		
Db	261	TGCGATTTCGTCGCGACATGTTTGGGCTCTGTTATTTAGAGCCTATGGCTTGACATCACTGA	320		
QY	819	ACAAGATTAACGTTTCAGAGATATGCACATGAACAAGTAAAGCACTCGTATGTTCGT	878		
Db	321	ACAGGATTGCAAAATTTGCAGAAAGATATGAGAGAGGAAGGATTGTCATCGTTGT	380		

QY	879	AAATTAAGGCGTACTGTGGGAAAAAGATGATGATAAAGCGATGGAAGAAATTTGAAGTGAAGT	938
DB	381	GACCAAGTGGGATGACGATCCCAACAGAACCAAGAACTACTGTATTATTATGAGAGA	440
QY	939	ACGTAA 945	
DB	441	TGTTGA 447	
RESULT 8			
BO281519			
LOCUS		532 bp	linear
DEFINITION	WHE3022_D01.G02Zs wheat unstressed seedling shoot normalized cDNA library Triticum aestivum cDNA clone WHE3022_D01.G02, mRNA		EST 13-MAY-2002
ACCESSION	BO281519		
VERSION	BO281519.1	GI:20549073	
KEYWORDS	EST.		
SOURCE	Triticum aestivum (bread wheat)		
ORGANISM	Triticum aestivum		
REFERENCE	Anderson O.D., Chao S., Choi D.W., Close T.J., Fenton R.D., Han P.S., Hsia C.C., Kang Y., Lazo G.R., Miller R., Nguyen H.T., Rausch C.J., Seaton C.B., Tong J.C. and Zhang D.		
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.		
	1 (bases 1 to 532)		
	Anderson O.D., Chao S., Choi D.W., Close T.J., Fenton R.D., Han P.S., Hsia C.C., Kang Y., Lazo G.R., Miller R., Nguyen H.T., Rausch C.J., Seaton C.B., Tong J.C. and Zhang D.		
	The structure and function of the expressed portion of the wheat genomes - Normalized shoot cDNA library		
	Unpublished		
JOURNAL	Contact: Olin Anderson		
COMMENT	US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel: 5105595773 Fax: 5105595818 Email: oanderson@pw.usda.gov Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: SK primer.		
FEATURES	Location/Qualifiers		
source	1..532		
	/organism="Triticum aestivum"		
	/mol_type="mRNA"		
	/cultivar="Chinese Spring"		
	/db_xref="taxon:4565"		
	/clone="WHE3022_D01.G02"		
	/tissue_type="Etiolated shoot"		
	/dev_stage="Five day old seedling"		
	/lab_host="E. coli DH10B"		
	/clone_lib="wheat unstressed seedling shoot normalized cDNA library"		
	/note="Vector: Lambda Uni-ZAP XR, excised phagemid plasmid SK; Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give plasmid phagemids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares'. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
BASE COUNT	162 a 112 c 110 g 148 t		
ORIGIN			
Query Match	7.6%; Score 99.2; DB 13; Length 532;		

Best Local Similarity 53.3%; Pred. No. 5.5e-10;
Matches 256; Conservative 0; Mismatches 218; Indels 6; Gaps 2

QY	753	TTTAAAGCCGTTAAACGTTCAAAATGGTTTAAAGCTTTATGCTTTATGTTGAGCAAGCAAGCAT	812
Db	48	TTTTCTTAAACATTTCTGGCCCTCTGATGTGGTTGCCCTTGTATCTTAGCCAAATGGCCCTGTGT	107
QY	813	TATTGAACAGATTAACGCTGTGGCAGATATGCACATGACAAAGGTAAAGCAGTGTGAT	872
Db	108	TACCGAGCAAGATTATATAAATTGACACAGAGATTGCACAAAGAGAAAGACTTGTGTGAT	167
QY	873	TGTCGTAATTAATGGGATCTACTGTGGAAAAAGATA--GTAAACGATGAGAAATTTTGA	929
Db	168	TGTTGTGACAAATGGGATCTACTATTCCTTAACACAGAACCATGAGATTAACACATTTATGA	227
QY	930	AGATGAAGTACGTTAAAGAAATTCGAATTTTATGATATATGCACAAATGCTTTTGTGTCTGC	989
Db	228	ACAGATGTATATAGAAAGCTTCCCATCTTGATTTGGGCACCATATTTGTCTACTGTCTTCTGC	287
QY	990	TAAAGAACGCACAAAGATTACGTACTTATTCCTCTTACATTAATGAAGCAAGTGAACCA	1043
Db	288	GACTTAATGTCCACAGTGTCAAAAGATTAATTTCTGCTCTCGTTTGGTGTGAGAAATGAAG	347
QY	1050	TAAAAAACGTTTCAAAAGTTCAACCTTAATTAAGAAGTCTTACTGATGTGCAATTTCCATGAA	1109
Db	348	GTCATTAAGACTCGGCACACTTCATTCTTAATCAAGTTATTTAGAGAAGCATATGCAATTCAA	407
QY	1110	CCCTACACCAACAGACAAAGGT--AGAAGTTTGAATCTTTTATATGCAACACAAGATTGC	1166
Db	408	ACCACACCAACAGAACAGAGGTGGCAAAAGAGGTGTCATATTATACAAACAGACGCTGC	467
QY	1167	TATGAACCAACCGCATTTTGTGTATTTGTTAATGTATGTAGCAATTAATGCATTTTCTTA	1227
Db	468	CATTGGTCCACCAACATTCGTTTATATTTATTAATATATGTCGAAGGCTCTTCCCTCGATACATA	527

RESULT 9	
A1465410/c	
LOCUS	
DEFINITION	A1465410 395 bp mRNA linear EST 09-MAR-1996 mus3bd2.x1 Stratagene mouse melanoma (#937312) Mus musculus cDNA clone IMAGE:552107 3' similar to SW:YPHC_BACSU P50743 HYPOTHETICAL 48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION., , mRNA sequence.
ACCESSION	A1465410
VERSION	A1465410.1 GI:4319440
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eumaxyloia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 395) Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Person Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurr,R., Ritzer ,E., Kohn,S.S., Shin,T., Jackson,T., Cardenas,M., McCann,R., Waterston,R. and Wilson,R. The WashU-NCI Mouse EST Project 1999 Unpublished
TITLE	
JOURNAL	
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999

```

FEATURES
source      Location/Qualifiers
1. 395      /organism="Mus musculus"

```

```

/ mol_type="mRNA"
/db_xref="taxon:10090"
/ clone="IMAGE:652107"
/ tissue_type="melanoma"
/ dev_stage="M2 cells"
/ lab_host="SOLR (kanamycin resistant)"
/ clone_lib="Stratagene mouse melanoma (#937312)"
/ note="Organ: skin; Vector: pluvscript SK; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. From M2 cells, a highly metastatic derivative of the
K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAAATCGCGCAGAG
3' -3' adaptor sequence: 5' CTCGACGTTTTTTTTTTTTTTT 3'"

```

Query Match	6.9%	Score 89.4	DB 9	Length 395
Best Local Similarity	55.3%	Pred. NO.5.3e-08		
Matches 194	Conservative	0	Mismatches 156	Indels 1
				Gaps 1

QY		955	TTTTAGACATTAATGCAAAATTGCCTTTGTGGTCCTTAAAGAACCCACAAGAATTACGTACA	1014
Dd		364	TTTTTAGCGTGAGCCCCAATTGTGTTTATTTCACG-ACACACAGGAAGTAGAATTTCITNAA	306
QY		1015	TTATTCCTTACATTAATGAAGCAGCAAGTGAANAACCATAAAAACGCTGTCCAAGTTCACCT	1074
Dd		305	TTAAGACAGCACCATTTTGTAGAAAGTCAAAAAATATATTAGATATAAAAAGTTTCACATATGCC	246
QY		1075	TTAAATGAAGTCGTTACTGATGCAATTTCCATGACCCCTACACCAACAGACAAAGTAGA	1134
Dd		245	CTTAATCAATTCATTTTAGATATATCAAAATGTTTACACACTCCCACAAGTGTCGTGAAAA	186
QY		1135	CGTTGAATGTCCTTTTATGCAACGCAAGTTCCTATAGACACCGACATTTGTTGTATTT	1194
Dd		185	AAATTAATATATTAAATTTGCTGCACCAAAATTCACCACAATATVTCCTATTTTATTTT	126
QY		1195	GTTAATGATAGTAGAATTAATGCACTTTTCTATTAAGCGCTATTATAGAAATCAATCCGT	1254
Dd		125	GTTAATGATATAAAATTTAATTCATTTTCTATCATAAAGATATATTTGAAATCAATTTAGA	66
QY		1255	GCCGCTTTTGGTTTGAAGTACCAACCAATTCATTTATAGCTGCAAGAGAGA	1305
Dd		65	AACATATTTTGGTTTTCACGTTGTCCAATTAAATTAATTTTATATAAAAAAAAA	15
RESULT 10				
CD377507				
LOCUS				
DEFINITION				
	CD377507	745 bp	mRNA	linear EST 31-May-2003
	PTM02676 Phaseodactylum tricornutum Uni-Zap XR Phaseodactylum			
	tricornutum cDNA 5' , mRNA sequence.			

RESULT 10	
LOCUS	CD377507
DEFINITION	CB377507 745 bp mRNA linear EST 31-MAY-2003
ACCESSION	PM0002676 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
VERSION	tricornutum cDNA 5', mRNA sequence.
KEYWORDS	CD377507
SOURCE	CD377507.1 GI:31253121
ORGANISM	EST.
	Phaeodactylum tricornutum
	Phaeodactylum tricornutum
	Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
	Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
REFERENCE	1 (bases 1 to 745)
AUTHORS	Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
TITLE	Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL	Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE	22111123
PUBMED	12114555
COMMENT	Contact: Bowler C Laboratory of Molecular Plant Biology Stazione Zoologica 'Anton Dohrn' Villa Comunale, I-80121, Napoli, Italy Tel: 39 081 583 3268/3211 Fax: 39 081 764 1355 Email: chrisealpha.szn.it Diatom EST database(http://aves.thagen.szbowler.com) Seq primer: T3 backward POLYA-No.

Query Match	6.4%	Score 83.2	DB 29	Length 1101
Best Local Similarity	19.6%	Pred. No. 1.1e-06		
Matches 130	Conservative 271	Mismatches 258	Indels 3	Gaps 1
BASE COUNT	201 a	64 c	131 g	202 t
ORIGIN				503 others
Query	441	ACATGTTTAGTCTTGGTACTGTTGATGACGTTGTTCTCATTTTGGTGAAGAGA	500	
Db	1099	ARRGGDDTWDHDTKKDMDWTWKMTWKDRADDRRMAGDADRMMWMDGAGTMTATWMMW	1040	
Qy	501	AGAGATCCCTATGATGACAGTACATTCGCTATTCATTTAGGACGACCAACGAGG	560	
Db	1039	MMWMTATDWTWKKMMWMTAAKTATTTATMTWTTTARADAGHDKAGKRDRAATDADAG	980	
Qy	561	TAAATCAAGTTTAAATGCTATTTTAAAGTGAAGATCGCTATTCGTTCTTAATGTTGC	620	
Db	979	RROGGRKRDKKKDDDDKGGKKKKAALAAWATKWMDDWMDKMDKAGAKRAD	920	
Qy	621	AGGACAACGAGACGCTATTTGATACAGATATGATTGATGACACAGATTATGTTT	680	
Db	919	DDDGAGRKDDGKKDDDDTDTGKDDDDKKKMDWMDKAKGTWGDATWMAATDMMWGW	860	
Qy	681	AATGATAGCTGCTGTTGCTTAAAGAAAGAAAGTATATCAATCACTGAAATATTC	740	
Db	859	ADADWMTWDAADWMAADDDRWDAAMKWDMAWGAFTADRDMDGRGKGGARKRRDR	800	
Qy	741	AGTATTAAGACCTTTAAAGCGATGAACGTTCAATGTTGTTTAAAGTATGATGC	800	
Db	799	KRADKRDADDDRDADATTTTWTTTTTR---DTDDWKKKIDTWTWRMAADRTWDDDDDR	743	
Qy	801	AGAACAAAGCATTTATTGAACAGATTAACGTTGTCAGAGATATGCACATGAACAGTAA	860	
Db	742	DRAGTAGKRWRTWKRWKRRDTWMDADADDTARDRRRRGDDGADAGKKGKTKGRKRR	683	
Qy	861	AGCAGTGGTGGTTTTCGTAATTAATATGCGATACGTGGAAGAAAGATAGTAACGATGAA	920	
Db	682	RDRATWDTWDAADWMAWMTTDTDTDDWDRDRRRRGARRRRRTTAAAWMDWTWRMD	623	
Qy	921	GAAATTTGAAGATGAAGTGAAGTGAAGTTCATTTTATGATTTATGACCAATTCGTT	980	
Db	622	WAKMDKTRADRDWRMAADTWDTRKADRDPAKRAMARARDDRAAARADRRWTKGTT	563	
Qy	981	TGTGCTGCTAAGACGACACAGATTAACGTATATTCCTTACATTAATGAAGCAAG	1040	
Db	562	TATWTTWMAARAAMWMAWMTATTTATWTTTWTWTTTWTWTTTWTWTTTAAWMAWTA	503	
Qy	1041	TGAAACACATTAAGAAAGCTGTTCAAAGTTCACGTTTAAATGAAGTCGTACTGATGCAT	1100	
Db	502	TWAAWTTAAWMAAAAAAATTTTWTTTTWTWMAWMAWMTATWTTTWTWTTTAAAT	443	
Qy	1101	TT 1102		
Db	442	TT 441		
RESULT 13				
LOCUS	BH234795	167 bp	DNA	linear
DEFINITION	MEAB_B03.y E Spiroplasma kunkelii genomic clone MEAB_B03.y, genomic			
ACCESSION	BH234795			
VERSION	BH234795.1			
KEYWORDS	GSS.			
SOURCE	Spiroplasma kunkelii			
ORGANISM	Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;			

REFERENCE	TITLE	AUTHORS	JOURNAL	COMMENT	
1	(bases 1 to 167)	Spiroplasmataceae: Spiroplasma.			
1	Hogenhout, S. A.	Genomic sequences from Spiroplasma kunkelii strain M2			
1	Department of Entomology				
1	The Ohio State University-OARDC				
1	120 Thorne Hall, 1680 Madison Avenue, Wooster, OH 44691, USA				
1	Tel: 330 263 3730				
1	Fax: 330 263 3686				
1	Email: hogenhout.1@osu.edu				
1	Plate: AA row: B column: 03				
1	Class: EcORI fragments				
1	Location/Qualifiers				
1	1. 167				
1	/organism="Spiroplasma kunkelii"				
1	/mol_type="genomic DNA"				
1	/strain="M2"				
1	/db_xref="taxon:47834"				
1	/clone="MEGA_B03.y"				
1	/clone_11b="E"				
1	BASE COUNT	58 a	26 c	26 g	57 t
1	ORIGIN				
1	Query Match	6.2%	Score 81.4	DB 28	Length 167
1	Best Local Similarity	68.7%	Pred. No. 2.1e-06		
1	Matches 112; Conservative 0;	Mismatches 51;	Indels 0;	Gaps 0;	
1	QY	533	TATCCATTATTGGACGACCAAACTAGCTAAATCACTTTAGTAAATGCTATTTAGGTG	592	
1	Db	5	TTTCTTTTAAATCGGTAACCAATGTTGGAAATCTTCTTAAACAAATGCTATTTAGGTG	64	
1	QY	593	AAGATCCGCTATGTTCTTATGTTTCAGGACAGACAGACGCTATTGATGATCAGAGT	652	
1	Db	65	AAGAACGAGTATTGTATTCACCAATTCGACGACAAACATCAGATTCATTCATCATCAT	124	
1	QY	653	ATAGTTATGATGACAGATTATGTTTAAATCGATCTGCTGG	695	
1	Db	125	TTAAACGAAATATCAACTTATCTGATTCGATTCATCTGCTGG	167	
1	RESULT 14				
1	LOCUS	B0855888	653 bp	mRNA	linear
1	DEFINITION	OG828A05.yg.ab1 OG:ABCDI lettuce salinas lactuca sativa cDNA clone			
1	ACCESSION	B0855888			
1	VERSION	B0855888.1	GI:22241353		
1	KEYWORDS	EST.			
1	SOURCE	Lactuca sativa			
1	ORGANISM	Lactuca sativa			
1	REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
1	AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
1		asterids; campanulids; Asterales; Asteraceae; Clctorioidae;			
1		Clochoriadeae; Lactuca.			
1		1 (bases 1 to 653)			
1		Kozik A., Michelmore R.W., Knapp S., Matvienko M., Rieseberg L.,			
1		Lih H., van Damme M., Lavelle D., Chevalier P., Ziegler J., Ellison			
1		P., Kolman J., Slabaugh M.S., Livingston K., Zhou Y., Val Z.,			
1		Church S., Jackson L. and Bradford K.			
1		Lettuce and Sunflower ESTs from the Composite Genome Project			
1		http://compgenomics.ucdavis.edu/			
1		unpublished			
1		Contact: Alexander Kozik (R.W. Michelmore)			
1		Department of Vegetable Crops, R.W. Michelmore Lab			
1		University of California at Davis (UCD)			
1		Asmundson Hall, UCD, Davis, CA 95616, USA			
1		Tel: 1-(530)-742-1742			
1		Fax: 1-(530)-752-9659			
1		Email: akozik@catc.org [michelmore@vegmail.ucdavis.edu]			
1		singleton, see http://cpdb.ucdavis.edu/ for details.			
1		Plate: OG828 row: A column: 05.			

FEATURES

Source

Location/Qualifiers
1. .653

BASE COUNT	200 a	111 c	156 g	186 t
ORIGIN				

Query Match	6.1%	Score 79.6;	DB 13;	Length 653;
Best Local Similarity	56.4%	Pred. No. 5.6e-06;		
Matches 195; Conservative	0;	Mismatches 139;	Indels 12;	Gaps 2

QY	485	ATTTTGGGAGAGAGAACATCTCTTTGATGGAAGATATACAAATTCGATATCCATTTTG	544
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	308	ATTCCATGTGGAAAACCAAGAGAGAAGAAAACACTACATTCCTTCAATTTCAATTTGAG	367
QY	545	GAGCACCAAACGTAGTAAATCAAGATTAGTAAATGCTATTTTAAAGTGAAGATCGCGTTA	604
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	368	GAACACCAAAATGTTGGAAAAAGCAGCATCTTCAAAAGCTTTAGTTGGTAAAGCACAACAA	427
QY	605	TCGTTTCAATGTTGACAGGACCAACGAGACGCGATTGATACAGAGT---ATAGTTATG	661
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	428	TTGTATAGTCCAGTATAGTGGAAACCAACGACTGACGCTATTTGACACTAGTTTGAAGACCA	487
QY	662	ATGACACAAGATTATGTTTTTAATCGACTGCTGGTATCGCTATATGAAAAAGAAAAAGTATATG	721
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	488	ATGGCCCAAAATTTGTTTGATTGATTAAGTATGCGTGGAAATTAAGAAAAAGGCACAGCAGTCTT	547
QY	722	AAT-----CAACTGAGAAATATTCAGTATTAAAGACCTTTAAAGACGATTTGAACGTT	772
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	548	CATCAGGAGATTAACAACAGAGCGTTTGTCTGTTAATCGAGCAATTCGCTGCAATTTGGCGTT	607
QY	773	CAATTCGTTTGTACTGCTTATGATGACAGAAACAAGCATTTATTTGA	818
	111	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111	
Db	608	CTGATGTGTAGCTCTCTGTTATTTAGAGGCTTTAGCACTGTATATCAACA	653

RESULT 15

LOCUS	501 bp	mRNA	linear	EST 05-JAN-1998
N37967				
DEFINITION	19194	Lambda-PRU2	Arabidopsis thaliana	CDNA clone 214C1777, mRNA

ACCESSION	N37967	
VERSION	N37967.1	GI:1159109
KEYWORDS	EST.	
SOURCE	Arabidopsis thaliana	(thale cress)
ORGANISM	Arabidopsis thaliana	

REFERENCE
1 (bases 1 to 501)
AUTHORS
Newman, T., deBruin, F. J., Green, P., Keegstra, K., Kende, H., McIntosh

TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)

MEDLINE

PUBMED

COMMENT

MEDLINE 95148729
PUBMED 7846151
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRR, Michigan State University, Plant Biology Bldg., E

FEATURES

source

Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers

BASE COUNT	136 a	92 c	135 g	118 t	20 others
ORIGIN					

Query Match	6.0%	Score 78.8;	DB 14;	Length 501;
Best Local Similarity	54.2%	Pred. No. 8e-06;		
Matches 219;	Conservative	0;	Mismatches 172;	Indels 13; Gaps 3

QY	499	GAACAAGATCCTTTATGATGAAGATGATCAATTCAGATTCATATCATATTTGGACGACCAACGTA	558
Db	17	GAAGGAGGGAAGAAAGAAACTACATTCCTCGCATTTGCAATTTATTTGGTGGCCAAATGTT	76
QY	559	GCTAAATCAAGTTTAAATGCTATTTTAAAGTGAAGATCGCGTTATCGTTTCTTAATGTT	618
Db	77	GGCAAAAGTACATTTTAAATGCACTTGCCGAGAGGATAGAAACATTTGTTAGCCCTGTT	136
QY	619	GCAGGGACACGAGAGACGCTATTGTATTCACAGTATA---GTTATGATGGACAAGATTAT	675
Db	137	AGTGGACATACCCCGATGCTATATGATCAGAAATTTACCGGACAGATGGGAGAAAGTTT	196
QY	676	GTTTAAATCGATACGCTGCTGATATGCGTAAAAAAGGAAAGATTAATGATCA-----	726
Db	197	CGGCTTATAGATACAGCTGGGATCAGGAAAAAGTATCTGTGGCTTCATCAGGACACT	256
QY	727	ACTGAGAATATTTCAGTATTTAAGACCTTTAAAGCGATTGAACGTTCAATGTGTTTTA	786
Db	257	ACAGAGGCGCATGTCACTGAATCGTGTGCATTTTCGAGCAATTCGTGTGTTGTGTTGC	316
QY	787	GTTGTTATTTGATGCGAACAAGGCATTTATTTGAACAAGATAAACGTGTTGCAGG-ATATGC	845
Db	317	CTTGTCATTTGAAGCTATGGCATGCTAAACAGAGCAGGATCTTAAAGNTTTCACGAGAGNATN	376
QY	846	ACATGACACAAGTAAGCAGCTGATTTGTCTTAATTAATGGG	889
Db	377	GNNANGAAGAGGNAAGGATGTCNTGTTGTGTNAACAATGGG	420

Search completed: September 12, 2003, 20:43:41
Job time : 3520.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 : Search time 5327.98 Seconds
(without alignments)
10020.134 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305

Sequence: 1 atgactaacactagtagc.....attatagctcgaagaga 1305

Scoring table: IDENTITY NUC

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_ov:*

5: gb_pat:*

6: gb_pl:*

7: gb_pr:*

8: gb_ro:*

9: gb_sts:*

10: gb_sy:*

11: gb_un:*

12: gb_vl:*

13: em_ba:*

14: em_fun:*

15: em_hum:*

16: em_in:*

17: em_mu:*

18: em_om:*

19: em_or:*

20: em_ov:*

21: em_pat:*

22: em_pl:*

23: em_ro:*

24: em_sts:*

25: em_un:*

26: em_vl:*

27: em_da:*

28: em_htg:*

29: em_in:*

30: em_mu:*

31: em_om:*

32: em_or:*

33: em_ov:*

34: em_pat:*

35: em_pl:*

36: em_ro:*

37: em_sts:*

38: em_un:*

39: em_vl:*

40: em_da:*

41: em_htg:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1292.2	99.0	301550	1 AP003134	AP003134 Staphyloc
2	1292.2	98.0	346900	1 AP003362	AP003362 Staphyloc
3	1290.6	98.9	1308	6 AX62268	AX62268 Sequence
4	1290.6	98.9	333750	1 AP004827	AP004827 Staphyloc
5	919.4	70.5	3269	1 AE270032	AE270032 Staphyloc
6	919.4	70.5	3269	6 AX145350	AX145350 Sequence
7	919.4	70.5	300892	1 AE016747	AE016747 Staphyloc
8	733.8	56.2	304680	1 AE017002	AE017002 Bacillus
9	727.4	55.7	290117	1 AE017028	AE017028 Bacillus
10	683.2	52.4	347050	1 AL591981	AL591981 Listeria
11	683.2	52.4	349980	6 AX641670	AX641670 Sequence
12	676.8	51.9	313450	1 AL596170	AL596170 Listeria
13	676.8	51.9	319630	6 AX413016	AX413016 Sequence
14	676.8	51.9	349980	6 AX417046	AX417046 Sequence
15	667.6	51.2	960	6 AX144037	AX144037 Sequence
16	656.8	50.3	300050	1 AP004599	AP004599 Oceanobac
17	647.4	49.6	24887	1 BACSERA	L47648 Bacillus su
18	647.4	49.6	213680	1 BSUB0012	Z99115 Bacillus su
19	634.8	48.6	302173	1 AE016951	AE016951 Enterococ
20	633	48.5	300550	1 AP001512	AP001512 Bacillus
21	600	46.0	1311	6 AX607165	AX607165 Sequence
22	600	46.0	20601	1 AE014265	AE014265 Streptoco
23	600	46.0	44145	6 AX602195	AX602195 Sequence
24	600	46.0	174050	1 SAG766852	AL766852 Streptoco
25	599.4	45.9	1311	6 AX433884	AX433884 Sequence
26	582.4	44.6	11071	1 AE006309	AE006309 Lactococc
27	552	42.3	52276	1 AE014141	AE014141 Streptoco
28	552	42.3	151947	2 SPNEU1902	AL449924 Streptoco
29	552	42.3	323825	1 AP005146	AP005146 Streptoco
30	550.4	42.2	1308	6 AX570314	AX570314 Sequence
31	550.4	42.2	5066	6 AR218847	AR218847 Sequence
32	550.4	42.2	5066	6 BD003759	BD003759 Polynucle
33	550.4	42.2	10310	1 AE007464	AE007464 Streptoco
34	550.4	42.2	349980	6 AX571764	AX571764 Sequence
35	550.4	42.2	349980	6 AX571765	AX571765 Sequence
36	549.6	42.1	12434	1 AE006498	AE006498 Streptoco
37	548.8	42.1	12370	1 AE009978	AE009978 Streptoco
38	548.8	42.1	12540	1 AE008523	AE008523 Streptoco
39	544.2	41.7	302050	1 AL935257	AL935257 Lactobaci
40	531.2	40.7	3737	1 AB016077	AB016077 Streptoco
41	531.2	40.7	13860	1 AE015016	AE015016 Streptoco
42	486.2	37.3	301278	1 AE015939	AE015939 Clostridi
43	476	36.5	296750	1 AP003191	AP003191 Clostridi
44	438.4	33.6	10861	1 AE007680	AE007680 Clostridi
45	423.2	32.4	10691	1 AE010530	AE010530 Fusobacte

ALIGNMENTS

RESULT 1
LOCUS AP003134/C
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS SOURCE
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE Staphylococcus aureus subsp. aureus N315
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Iian, J., Ito, T., Kanamori, M.,
Matsunaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

TITLE Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C., Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J., Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C., Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*

JOURNAL Lancet. 357 (9264), 1225-1240 (2001)

MEDLINE 21311952

REFERENCE 2 (bases 1 to 301550)

AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.

TITLE Direct Submission

JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome-49-10 Nishihara, Shibuya-Ku, Tokyo 151-0066, Japan (E-mail:biocente.go.jp, URL:http://www.btc.nite.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424)

COMMENT On Jun 12, 2001 this sequence version replaced gi:13701258.

FEATURES

source

1. 301550

Location/Qualifiers

organism="Staphylococcus aureus subsp. aureus N315"

mol_type="genomic DNA"

strain="N315"

sub_species="aureus N315"

db_xref="taxon:15879"

complement(249..947)

gene="SA1293"

complement(249..947)

gene="SA1293"

note="ORFID:SA1293"

codon_start=1

transl_table=11

product="conserved hypothetical protein"

protein_id="BAB42554.1"

db_xref="GI:13701259"

translation="MSIFTMITIVYILMLPLPAHQKVSNEYKYSQVRSCTGR

EVAFILIHAGLYDEVVKGEGFLTDHDPKRVVLSLPANDRPSVAGTAIAHEVG

HAIDHQGYWFLFRALVPAVNLGSSLSYIMLGIITAIAGSAGSTALMIGLM

SLAVFLSIVTLPEFEDASSRAMKQITALINVEKEYKARKVLSAAMTYAATVAV

ALVRLIILARSDD"

complement(1001..1588)

gene="SA1294"

complement(1001..1588)

gene="SA1294"

note="ORFID:SA1294"

codon_start=1

transl_table=11

product="conserved hypothetical protein"

protein_id="BAB42554.1"

db_xref="GI:13701260"

translation="MTINTFMOYLYLQORSWMLMLLIGNLGIYGYIWMGEOLSHTP

HKIFVDPSPATILYVLSILTIKONSIIEALAPTLFEYGIAMVIMNLFTIEO

GDIYNGLVMSHSIMAVOAIFYPRFRSMIGISVAMIVFLNDYIDYFIHQPPY

DFITHWQIGVLSCLSFGLLTIELKLKCK"

complement(1578..2153)

gene="SA1295"

complement(1578..2153)

gene="SA1295"

note="ORFID:SA1295"

codon_start=1

transl_table=11

product="conserved hypothetical protein"

protein_id="BAB42554.1"

db_xref="GI:13701261"

translation="MSEITNQKSEFIEYLLFOYRFSRIAMVNIKYNKAKIANI

HFVDKINHHETLEAEVGSASATQTRNKIKMLNTEIFDIYANHHCAFDIQTHFA

NVSKREORLDDILVAQTESPSYQYLDLNSMALIDRKHALLDYLLAHNIDLSQMN

EKORFYQLTOILNLTKLKLVKHNQFEDLADDD"

complement(2167..3411)

gene="SA1296"

complement(2167..3411)

gene="SA1296"

note="ORFID:SA1296"

codon_start=1

transl_table=11

note="ORFID:SA1296"

codon_start=1

transl_table=11

product="conserved hypothetical protein"

protein_id="BAB42556.1"

db_xref="GI:13701262"

translation="MEDYIKLIDIDINLQKLEMLDSRVNAITTDNDALFILGETLYN

FGMLPQGLEFVRVLYHKYPDESELLIYFIEGLMSERQKDEALEYLSYVEPSEKMLE

ADLYOINMMEVAILDLQALELEPNDDPIHFALAMELYDQYLRATSEYETLETG

EYOVANGVNLFSRMADCSLOSQNSDAIRLYDENINDEMETSEDLKKAISYEDNITQ

AKIKMTTSLKDPDYTOGYLYLSIYENKENYPDALIEGKEGLRSOFYKELMYTGC

LEIEHGDNESVLLKQALEVNAQOEPLILISDLYRNEDEYEAITTELTLYDEEDID

PTFMHIAFATQGEERDKAQHFELAIPTKNTINDPSDYIFYLIEIGQKQAITTL

NOLLEPSENMWHDLSRLQS"

complement(3418..4716)

gene="aroA"

complement(3418..4716)

gene="aroA"

note="ORFID:SA1297"

codon_start=1

transl_table=11

product="3-phosphohikimate 1-carboxyvinyltransferase"

protein_id="BAB42557.1"

db_xref="GI:13701263"

translation="WVSEQIIDISGPLAGEIEVPGDKSMTIRAIMLASIGTSNIYK

PLLGEDCRRTMDIFRLAGVDIKEDDKLVNPSGYKAKRTPQVLYTGSGTTRBLA

GILSGIGIESVLDGVSIGKRPMDHVLRPLKMDANIIEGIDENYPLIIEPSVIGIN

YOMEVASAOVKSAILEASLFSNDTVYIKELDVSRNHTEFMFHMPIRSEELSTTT

PDALIOHKPADHPVGDISSAFAFIYAALITPESDVTIHNQINFTPSGELIIDIVKMG

GNIOIFNOTGAEPTASIRIOTYTPMLQPTIEGELVPRKADLPVIALICQAVCTST

IKDAELKAKETNRIDTTRADMLNLGFELOPTINDCLIIHPSEFKTNATVYDSLTDHRIG

KMLAVASLSSSEPVKIKOPDAVNSFPGLPLKLEENEG"

complement(4726..5790)

gene="aroB"

complement(4726..5790)

gene="aroB"

note="ORFID:SA1298"

codon_start=1

transl_table=11

product="3-dehydroquinate synthase"

protein_id="BAB42558.1"

db_xref="GI:13701264"

translation="MKLOTTPSSNNYPFIVEGALDHSITYIDQFQSFILDEHVNQ

YFADKFDDLSTENHAKVITIPAGEKTKPEOYQETLEYILSHVTRNAILAVGGAT

GDFAGFVAATILRGVHFIOVPTTILAHDSVGKVGINSKQKRLIGAFRPTAVIYD

IDFLKTLPEQILSGYAEVYKHALNGESTQDEIQHKRDKREILQSLGMDKYIAKGI

ETKLDIVADEKEGCVRRFLNGLGTFEGHAEVYHNKIGASHAVMIGITIOFYANILFN

SNHDIQHYINVTYLRGYPLETTIDDEFIYQYMLSKKNDQGVWLIKRFQDIIV

QHIDQTTQAHACEQLKTYFK"

complement(5816..6982)

gene="aroC"

complement(5816..6982)

gene="aroC"

note="ORFID:SA1299"

codon_start=1

transl_table=11

product="chorismate synthase"

protein_id="BAB42559.1"

db_xref="GI:13701265"

translation="MRYLTSGESHGSPOLYIVEGIPANLEIKVEDINKEMFRQGGYG

KGRMQIEKDYIEYSGVRNGTTLSPITTYVYTNDDFTHMKRIKGAAPISSEERNNK

RTTYPRGHADLVGMYKYNHRDLNRVLEERSARETAKRAVAAGALCKVLLQDLIDY

SVVEIGIKDKDQDSETFKANLDNRVRIYDSDIAQMRBKIDEAKNEGSIGGV

GVVVENMPYGVGSVYHDKRLDGKIAQGVSVINARKGVSGFGFAARKPGSEIODEI

LYNSEIYVGRSGNHGLEGSGNMPITVNGVMKPIPLYKPLNSVDINTKDEKAT

IERSSCAVPASIVCEHYVAFEIKALLEEQSHHIDLOQIADRQOLANVE"

complement(7368..7568)

gene="SA1300"

complement(7368..7568)

gene="SA1300"

note="ORFID:SA1300"

codon_start=1

transl_table=11

```

/product="hypothetical protein"
/protein_id="BAB42560.1"
/db_xref="GI:13701266"
/translation="MKNISILGILMLKMLLYGAIFLDKPLNDITKIIIFILFLI
YIFVYLKELIIFLKSINKSKSA"
complement(7597..7722)
/gene="SAS045"
complement(7597..7722)
/gene="SAS045"
/feature="ORFID:SAS045"
/feature="ORFID:SAS045"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42561.1"
/db_xref="GI:13701267"
/translation="MWLIMITHNRARFLCCSKLLEKPFYFLVIRIKRGIFIK"
complement(7777..8226)
/gene="ndk"
complement(7777..8226)
/gene="ndk"
/feature="ORFID:SA1301"
/feature="ORFID:SA1301"
/codon_start=1
/transl_table=11
/product="nucleoside diphosphate kinase"
/protein_id="BAB42562.1"
/db_xref="GI:13701268"
/translation="MERFLLKPKPAVONRLIGEVISRIERKGLKLVGCKLMQPMEL
AETHYGEHQGRPFYNDLSIFISAPVFAVYVGEGDAVNSRHIGSTNPSEASPGSIR
GDGLGTIVRQIHGSDLSIESAREINLMENENETSYASPRDAMLYE"
complement(8318..9277)
/gene="gerCC"
complement(8318..9277)
/gene="gerCC"
/feature="ORFID:SA1302"
/feature="ORFID:SA1302"
/codon_start=1
/transl_table=11
/product="heptaprenyl diphosphate synthase component II"
/protein_id="BAB42563.1"
/db_xref="GI:13701269"
/translation="WAKLNMMNEIKRVEGRLEKAIKSKDVSLEQASLHLISGGKRYR
PAFVLSQFKEDEKTSQTYVAALGLIHNAATLVHDDVIDKSDKRGKLTYSKKWD

```

Query Match	99.0%	Score 1292.2	DB 1:	Length 301550:
Best Local Similarity	99.4%	Prod.No. 1.3e-214:		
Matches 1297:	Conservative	0:	Mismatches 8:	Indels 0: Gaps 0:
QY 1	ATGACCTAAACCTATTAGTACCTATTGTAGTGAGCCCTAAATGTAGTAAATCTCAATTTT	60		
Db 13778	ATGACCTAAACCTATTAGTACCTATTGTAGTGAGCCCTAAATGTAGTAAATCTCAATTTT	13719		
QY 61	AATGAAATAGTGGAGAACGTGTTTCGATTGTTGAAGATACGCCAGGTAAACACAGAT	120		
Db 13718	AATGAAATAGTGGAGAACGTGTTTCGATTGTTGAAGATACGCCAGGTAAACACAGAT	13659		
QY 121	CGATTATATTCCTCAGGTAGTGGTTAAACAATGATTCAATATTAATTATGATACAGTGT	180		
Db 13658	CGATTATATTCCTCAGGTAGTGGTTAAACAATGATTCAATATTAATTATGATACAGTGT	13599		
QY 181	ATTGAAATTTGGTATGCACCATTTCCAAACACAATATTAGAGCCAGACAGAAATGCCATA	240		
Db 13598	ATTGAAATTTGGTATGCACCATTTCCAAACACAATATTAGAGCGAGCAAAATGCCATA	13539		
QY 241	GATGAACCGGATGTTATTATTATTTTATGTTAACGTGCGTGAAGGATTGACACAAAGCGAT	300		
Db 13538	GATGAACCGGATGTTATTATTATTTTATGTTAACGTGCGTGAAGGATTGACACAAAGCGAT	13479		
QY 301	GAATGTGCTCCCTCAAAATTTTATACAAATCTAAAAAACCAGTGTATTACGGTTAAACAA	360		
Db 13478	GAATGTGCTCCCTCAAAATTTTATACAAATCTAAAAAACCAGTGTATTACGGTTAAACAA	13419		
QY 361	GTGATATATATTCGAAATGCGTACAGACGCTGATGATTTCTATTCATTAGATTGGTGA	420		
Db 13418	GTGATATATATTCGAAATGCGTACAGACGCTGATGATTTCTATTCATTAGATTGGTGA	13359		

OY	421	CCGATATCCAAATATACAGGTCTCATCATGTTTACGTTTACGTTGACCTGTTAAGATGCAGTTGTT	480
Db	13358	CCGATATCCGAATATAGGGTGCACATGTTTACGTTTACGTTTACGTTGACCTGTTAAGATGCAGTTGTT	1329
OY	481	TCTCATATTTGGTGAAGAGAGAAAGATCCCTATGATGATGAAGATACAAATTCGACTATCCATT	540
Db	13298	TCTCATATTTGGTGAAGAGAGAAAGATCCCTATGATGATGAAGATACAAATTCGACTATCCATT	1323
OY	541	ATTGAGACACAAACGTAAGTAAATCAAGTTTAAATGCTATTTTAAAGTGAAGATGCG	600
Db	13238	ATTGAGACACAAACGTAAGTAAATCAAGTTTAAATGCTATTTTAAAGTGAAGATGCG	1317
OY	601	GTTATTCGTTTCTAATGTTTGCAGGGACAAACGAGACCCCTATGATACAGATATAGTTAT	660
Db	13178	GTTATTCGTTTCTAATGTTTGCAGGGACAAACGAGACCCCTATGATACAGATATAGTTAT	1311
OY	661	GATGAGACAAAGATATATGTTTAAATGATACCTGCTGATGCTGTAATAAAAAAGAAAGATATAT	720
Db	13118	GATGAGACAAAGATATATGTTTAAATGATACCTGCTGATGCTGTAATAAAAAAGAAAGATATAT	1305
OY	721	GAATCAACTGAGAAATATTCAGTATTAAGACGTTTAAAGCGATTGAACGTTCAATGTT	780
Db	13058	GAATCAACTGAGAAATATTCAGTATTAAGACGTTTAAAGCGATTGAACGTTCAATGTT	1299
OY	781	GTTTATAGTGGTATATGATGCGAACAAGGCATTTATGAACAAGTAAACGTTGTCAGCA	840
Db	12998	GTTTATAGTGGTATATGATGCGAACAAGGCATTTATGAACAAGTAAACGTTGTCAGCA	1293
OY	841	TATGCACATGAACAAGGTAAAGACAGTGGTGTTCTGTAATAATAGGATGACTGTGAA	900
Db	12938	TATGCACATGAACAAGGTAAAGACAGTGGTGTTCTGTAATAATAGGATGACTGTGAA	1287
OY	901	AAAGATAGTAAACGATGAGAAAGATTTGAAGATGAAGTACGTAAAGATTTCCAAATTTTGA	960
Db	12878	AAAGATAGTAAACGATGAGAAAGATTTGAAGATGAAGTACGTAAAGATTTCCAAATTTTGA	1281
OY	961	GATTATATGACAAATTTGCTTTTGTCTGTGCTTAAGAACGACACATTAACGTACATTTTTC	1020
Db	12818	GATTATATGACAAATTTGCTTTTGTCTGTGCTTAAGAACGACACATTAACGTACATTTTTC	1275
OY	1021	CCTTACATATTAATGAAGCAAGTGAACACCATTAATAAAACGTTGTTCAAGTTCAACTTTAAAT	1080
Db	12758	CCTTACATATTAATGAAGCAAGTGAACACCATTAATAAAACGTTGTTCAAGTTTAAAT	1269
OY	1081	GAAGTCGTTACTGATGCAATTTCCATGATACCCCTACACCAACAGACAAAGGTAGACGTTTG	1140
Db	12698	GAAGTCGTTACTGATGCAATTTCCATGATACCCCTACACCAACAGACAAAGGTAGACGTTTG	1263
OY	1141	AATGTCGTTTATGACACAAAGTTGCTATGAACACACGACATTTGTTGATTTGTTAAAT	1200
Db	12638	AATGTCGTTTATGACACAAAGTTGCTATGAACACACGACATTTGTTGATTTGTTAAAT	1257
OY	1201	GATGTAGAAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAAAATCCGTCGCCCT	1260
Db	12578	GATGTAGAAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAAAATCCGTCGCCCT	1251
OY	1261	TTTGCGTTTGAAGGTACACCAATTCATATTTATAGCTGGAAGAGCA	1305
Db	12518	TTTGCGTTTGAAGGTACACCAATTCATATTTATAGCTGGAAGAGCA	12474
RESULT 2			
AP003362/c			
LOCUS	AP003362	346900 bp	DNA
DEFINITION	Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete sequence, section 5/9.		
ACCESSION	AP003362 BA000017		
VERSION	AP003362.2 GI:14247083		
KEYWORDS	Staphylococcus aureus subsp. aureus Mu50		
SOURCE	Staphylococcus aureus subsp. aureus Mu50		
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.		

```
REFERENCE
AUTHORS      1
              Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
              Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
              Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
              Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
              Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
              Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
              Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
              Whole genome sequencing of methicillin-resistant Staphylococcus
              aureus
JOURNAL      Lancet 357 (9264), 1225-1240 (2001)
MEDLINE      21311952
PUBMED       11418146
REFERENCE     2 (bases 1 to 346900)
AUTHORS      Ohta,T.
TITLE         Direct Submission
              Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
              of Medical Technology and Nursing, Department of Medical
              Technology, 1-1-1 Ten-nooda, Tsukuba, Ibaraki 305-8577, Japan
              (E-mail:tohts@tsukuba.ac.jp, Tel:81-298-53-3454,
              Fax:81-298-53-3454)
              On May 29, 2001 this sequence version replaced gi:13875626.
COMMENT
FEATURES
source       1. 346900
              /organism="Staphylococcus aureus subsp. aureus Mu50"
              /mol_type="genomic DNA"
              /strain="Mu50"
              /sub_species="aureus Mu50"
              /db_xref="taxon:158878"
              12. .263
              /gene="SAV1313"
              12. .263
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
              /protein_id="BAB57475.1"
              /db_xref="GI:14247084"
              /translation="MLTLIKLEDEQVILIEYIPEDDISNGKSVTFNNKDAEVIDFS
              LSEINEEYFMILYRNKSFVDRPFLEKQEPENYRIAWY"
              987. .1181
              /gene="SAV1314"
              987. .1181
              /gene="SAV1314"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
              /protein_id="BAB57476.1"
              /db_xref="GI:14247085"
              /translation="MNLQSYIGKLVKLTLTNNKILIGKVIDEDDKVDNFDGYNSIED
              TGRITVDISENKKIRIVLQK"
              complement(1725. .2750)
              /gene="SAV1315"
              complement(1725. .2750)
              /gene="SAV1315"
              /codon_start=1
              /transl_table=11
              /product="conserved hypothetical protein"
              /protein_id="BAB57477.1"
              /db_xref="GI:14247086"
              /translation="MISFENDILEGAHEKVLKRLVDTNLVQASGYGFDOFTQALIEKI
              KTDIDCPNATIEFLVGQTQNOVINSIMLESIEGYISADTGVAHVEGGALEYSHKV
              ITIPSEKGVASDVEYMETFEKSPDKDMVPGMVIYISHTPEGYTLYSKSELEELK
              KVQCYQPLPFMDGARGGLMSQSDMDITIDIAIKYCVFYIGTKIGALGCEALVFT
              KNNPKQFTTRIKHKGALLAGRLGTIOLELFTNLEYNISRHAIEAMNKKKDGFEKN
              KGYRLYDPSPTNQOFFILSNEXIALLEQKVFKAWEKIDDOHRYVRPATSWATTEENL
              NKLEELT"
              3020. .3217
              /gene="SAV1316"
              3020. .3217
              /gene="SAV1316"
              /codon_start=1
              /transl_table=11

gene
CDS          /product="hypothetical protein"
              /protein_id="BAB57478.1"
              /db_xref="GI:14247087"
              /translation="MEQIKTKFTAEELLEKNINAFLSSEATNLKLVNITKEIE
              ERFPNNEEPNALILLSVNK"
              3238. .4755
              /gene="SAV1317"
              3238. .4755
              /gene="SAV1317"
              /codon_start=1
              /transl_table=11
              /product="cardiolipin synthetase homolog"
              /protein_id="BAB57479.1"
              /db_xref="GI:14247088"
              /translation="MEFYPPFIEGKTMRYTFESNDIGTLFTIILAIGFILNVLAFIII
              FLERNRRATSTWAMLEVLVPLIGFILYLFEGRTVSARKLNKNGNVLTPEDGLK
              QOIESPDKNGVGTGDKNOKHNDIVSMILMDQDGTLENNKVDHFDGNDLYDQVLDK
              INAKETIILEYTRALDGLGRILHLEERKQGLPYKILYDDVGSRYKVNANDHF
              KSLDGEVEAFKSKLPFLNFRNNRNHRIIVYDQGLGVGFNIGDELGLGKLGIV
              RDTHLRIQGDVAVALQRLFLDMNSQAHRPQEPDYVKPRKNGPLGNSPIQIIAASGP
              ASDWQIEGYGPKIMISAKSVYLQSPYIIPNSYINAIKIAKSGVDVHLMIPEKPD
              HDLVYWPATPSNADLSGKVIYTYENGFIHSKMCILIDEIVSVGTANMPFSPELNF
              EVNAFYDENLAKDLRVANENDITTSKQLTRESYANRPLSVAFKESLAKVSPIL"
              4939. .5436
              /gene="SAV1318"
              4939. .5436
              /gene="SAV1318"
              /note="ATP-binding protein"
              /codon_start=1
              /transl_table=11
              /product="ABC transporter homolog"
              /protein_id="BAB57480.1"
              /db_xref="GI:14247089"
              /translation="MIOISNINKNSFNKRCVLKNISFDEQGCIALIGKNGAKSTLI
              DILGVNANSGEIFPDKDLQSENRSIMFQKTPMDLKVETIINLQSPFENPLPL
              ERITLTKRDSQNLQNFVKLSGGQRLDLPVLSLIGQQLILDEPSTNMIDIREY
              FMSIT"
              5461. .5838
              /gene="SAV1319"
              5461. .5838
              /gene="SAV1319"
              /note="ATP-binding protein"
              /codon_start=1
              /transl_table=11
              /product="ABC transporter homolog"
              /protein_id="BAB57481.1"
              /db_xref="GI:14247090"
              /translation="MLYTSHYIEEVRMSDKIILLENGLIINDSTSHRTNQSQIT
              LSEYTRKLKLDKDDLVITQKNHNGTIIITTSNVNDTILYLQOLHNLDDIEIYQVSI
              DSYFNNKKQRGSNYDTKLEENRI"
              5807. .6538
              /gene="SAV1320"
              5807. .6538
              /gene="SAV1320"
              /note="similar to ABC transporter integral membrane
              protein"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
              /protein_id="BAB57482.1"
              /db_xref="GI:14247091"
              /translation="MILSYLKEFVIMNKKTTLLISILFPVIFYILFTSILPEDEV
              KRFYKREYMSWTNVLISLFSLTPLDIINKQEMQORLWNPFPFTSYISYKVV
              TMLQPAIILIVLFVNGHYKGVAMSAVOMLESGIFLMGASLLIFGLIFSLNDIQ
              TSAIANVITGLIVAGGIMRPIINPNNLQHAHVLPSYHLRKLKGVDLASNNHNLIS
              FAIILLYLGSITIAVYCISHRRRAE"
              6542. .7633
              /gene="SAV1321"
              6542. .7633
              /gene="SAV1321"
              /note="similar to two-component sensor histidine kinase"
              /codon_start=1
              /transl_table=11
```

```
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKFLKDTSAIEISLILYLPINAGIFENVEYVGGPKMLIISVIYF
SLSYLIVYNNRNLNTIMEYFILILIHFIITCYEFSVHPMLSEFFSARAVPTFEK
NVRKATNLEILTMICTIITTYLLYNNFYAMMYVYVILMLDNKKNNRKYQKE
IAKNRHINTLIAEOERHRIGODHLHFAVSLISKESEAYKLIDTDEKVAELL
AINKLSRESINIKREIIDVDKIDPSFIEIDISIRKLDADIDFTEENKELAOVLSPTK
OSMLVMTTREAINNVIKHANASKVHGKLVNNHKLMLIEDDCKIDSDCEVKSISO
RNOHNGTAVDSTNGCKITIEISTGIA"
7630..8232
/gene="SAV1322"
7630..8232
/feature="SAV1322"
/feature="similar to two-component response regulator"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MTSLIAEDONMLROAMVOLIKLHGDEILADPDNGIDAMKLE
EYNNVYIIDEMPGMGLEVAIEIRKKHLNIIKIYITTKRREGYFEKAVYNDIAY
LKERSIEELVEITIKNVNKEKESATILTSFVDKNPLTPEQOVLKREIGLSSKEI
SEKLFITDGTIVRNTSVIIDLFPADNRFDMKKRNEKGWT"
complement(8345..8533)
/gene="SAV1323"
complement(8345..8533)
/gene="SAV1323"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNTSMENTVTLIAVIAIIIGITLQIFKPLIVTAVLSILLGI
FVGRTVLIVSFKRKN"
8672..9205
/gene="nuc"
8672..9205
/gene="nuc"
/feature="SAV1324"
/codon_start=1
/transl_table=1
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"

Query Match          99.08; Score 1292.2; DB 1; Length 346900;
Best Local Similarity 99.48; Pred. No. 1.2e-214;
Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

gene
CDS
61
200810
200750
200690
200630
200570
200510
200451
```

```
QY 361 GTAGATATATGAAATCGCTACAGACGTGTATGATTTCTATTCATTAGATTTGGTCAA 420
DB 200450 GTACATATATGAAATCGCTACAGACGTGTATGATTTCTATTCATTAGATTTGGTCAA 200391
QY 421 CCGTATCCCATATATCAGGCTCATGCTTTAGCTTGGTGACTTGTAGATGACGTTGTT 480
DB 200390 CCGTATCCCATATATCAGGCTCATGCTTTAGCTTGGTGACTTGTAGATGACGTTGTT 200331
QY 481 TCTCATTTTGGTGAAGGAAAGATCCCTATATGATGAGTATACAAATTCGCTATCCATT 540
DB 200330 TCTCATTTTGGTGAAGGAAAGATCCCTATATGATGAGTATACAAATTCGCTATCCATT 200271
QY 541 ATTGACGACCAACAGTATGATTAATCAAGTTTATTAATGCTATTTTAGTGAAGATCGC 600
DB 200270 ATTGACGACCAACAGTATGATTAATCAAGTTTATTAATGCTATTTTAGTGAAGATCGC 200211
QY 601 GTTATCGTTTCTTAATGTTGACGAGCAACGAGACGCTATTTGATACAGACTATATGTTAT 660
DB 200210 GTTATCGTTTCTTAATGTTGACGAGCAACGAGACGCTATTTGATACAGACTATATGTTAT 200151
QY 661 GATGACAAAGATTTATGTTTAAATGATATCTGCTGATGCTGATGCTGATGCTGATGCTGAT 720
DB 200150 GATGACAAAGATTTATGTTTAAATGATATCTGCTGATGCTGATGCTGATGCTGATGCTGAT 200091
QY 721 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGGATTTGAACGTTCAAAATGTT 780
DB 200090 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGGATTTGAACGTTCAAAATGTT 200031
QY 781 GTTTTATGCTTATTTATGATGACGAAACAGCATTTATGAACAGATTAACGTTGTCACGA 840
DB 200030 GTTTTATGCTTATTTATGATGACGAAACAGCATTTATGAACAGATTTAAGGATTTGTCACGA 199971
QY 841 TATGACATTTGAACAAGTAAAGCACTGCTGATTTGCTGAATTAATGGGATACGTGGAA 900
DB 199970 TATGACATTTGAACAAGTAAAGCACTGCTGATTTGCTGAATTAATGGGATACGTGGAA 199911
QY 901 AAAGATAGTAAAGATGAAGAAATTTGAAATGATGATGATGATGATGATGATGATGATGAT 960
DB 199910 AAAGATAGTAAAGATGAAGAAATTTGAAATGATGATGATGATGATGATGATGATGATGAT 199851
QY 961 GATTATGCAAAATTTGCTTTTGTGCTGCTGAAGAACGACACAGATTAAGATTAATGATTAAT 1020
DB 199850 GATTATGCAAAATTTGCTTTTGTGCTGCTGAAGAACGACACAGATTAAGATTAATGATTAAT 199791
QY 1021 CCTTACATTTATGAAGCAAGTGAAGAAACCATTAAGAAACCTGTTCAAGTTCAACTTAAT 1080
DB 199790 CCTTACATTTATGAAGCAAGTGAAGAAACCATTAAGAAACCTGTTCAAGTTCAACTTAAT 199731
QY 1081 GAAGCTGTACTGATGCAATTTCCATGAACCTTACACCAAGCAACAGCAAGATTAAGATTAAT 1140
DB 199730 GAAGCTGTACTGATGCAATTTCCATGAACCTTACACCAAGCAACAGCAAGATTAAGATTAAT 199671
QY 1201 GATGATGAATTAATTCATTTTCTTATTAAGCTTATTAAGCAATCAATCCGTCGCT 1260
DB 199610 GATGATGAATTAATTCATTTTCTTATTAAGCTTATTAAGCAATCAATCCGTCGCT 199551
QY 1261 TTTGGTTTGAAGTACACCAATTCATTTATTAAGCTGCAAGAGAGA 1305
DB 199550 TTTGGTTTGAAGTACACCAATTCATTTATTAAGCTGCAAGAGAGA 199506

RESULT 3
AX622668 1308 bp. DNA linear PAT 20-FEB-2003
LOCUS AX622668
DEFINITION Sequence 5631 from Patent WO02094868.
ACCESSION AX622668
VERSION AX622668.1 GI:28450653
KEYWORDS
```

```

SOURCE      Staphylococcus aureus
ORGANISM    Staphylococcus aureus
REFERENCE   1
AUTHORS     Masignani,V.C., Mora,M.C. and Scarselli,M.C.
TITLE       Staphylococcus aureus proteins and nucleic acids
JOURNAL     Patent: WO 02094868-A 3631 28-NOV-2002;
            Chilton Spa (IT)
FEATURES
  source
    1..1308
    /organism="Staphylococcus aureus"
    /mol_type="genomic DNA"
    /db_xref="taxon:1280"
BASE COUNT  450 a      184 c      278 g      396 t
ORIGIN
Query Match      98.9%  Score 1290.6; DB 6; Length 1308;
Best Local Similarity 99.3%; Pred. No.9.9e-214;
Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTACTATTTGAGGTAGGCGCTTAATGTAATCTACAATTTT
    |||||
Db 1 ATGACTAAACCTATAGTACTATTTGAGGTAGGCGCTTAATGTAATCTACAATTTT
    |||||
QY 61 AATGACATAGTTGGAGAACGCTGTTGCAATGTTGAAGATACGCCAGCTTAACAGAGAT
    |||||
Db 61 AATGACATAGTTGGAGAACGCTGTTGCAATGTTGAAGATACGCCAGCTTAACAGAGAT
    |||||
QY 121 CGATTTATTTCTTCAGGTGAGTGAATACATGATTTCAATATTTATGATCAGGTGT
    |||||
Db 121 CGATTTATTTCTTCAGGTGAGTGAATACATGATTTCAATATTTATGATCAGGTGT
    |||||
QY 181 AATTGAATTTGGTATGACACCATTTCCAAACAAATTTAGAGCGAGCAGAAATCCGCATA
    |||||
Db 181 AATTGAATTTGGTATGACACCATTTCCAAACAAATTTAGAGCGAGCAGAAATCCGCATA
    |||||
QY 181 AATTGAATTTGGTATGACACCATTTCCAAACAAATTTAGAGCGAGCAGAAATCCGCATA
    |||||
Db 181 AATTGAATTTGGTATGACACCATTTCCAAACAAATTTAGAGCGAGCAGAAATCCGCATA
    |||||
QY 241 GATGAACGCGATGTTATTTATTTATGTTAAACGTCGTTGAAGATTTGACAAACGAT
    |||||
Db 241 GATGAACGCGATGTTATTTATTTATGTTAAACGTCGTTGAAGATTTGACAAACGAT
    |||||
QY 241 GATGAACGCGATGTTATTTATTTATGTTAAACGTCGTTGAAGATTTGACAAACGAT
    |||||
Db 241 GATGAACGCGATGTTATTTATTTATGTTAAACGTCGTTGAAGATTTGACAAACGAT
    |||||
QY 301 GAAATGTCGTCCTCAATTTTATACAAATGTAACAAACCGTCGTTAGCGTTAACAA
    |||||
Db 301 GAAATGTCGTCCTCAATTTTATACAAATGTAACAAACCGTCGTTAGCGTTAACAA
    |||||
QY 301 GAAATGTCGTCCTCAATTTTATACAAATGTAACAAACCGTCGTTAGCGTTAACAA
    |||||
Db 301 GAAATGTCGTCCTCAATTTTATACAAATGTAACAAACCGTCGTTAGCGTTAACAA
    |||||
QY 361 GTAGATTAATATGAAATGCGTACAGACGCTGATGATTTCTATTCATTTAGATTTG
    |||||
Db 361 GTAGATTAATATGAAATGCGTACAGACGCTGATGATTTCTATTCATTTAGATTTG
    |||||
QY 361 GTAGATTAATATGAAATGCGTACAGACGCTGATGATTTCTATTCATTTAGATTTG
    |||||
Db 361 GTAGATTAATATGAAATGCGTACAGACGCTGATGATTTCTATTCATTTAGATTTG
    |||||
QY 421 CCGTATCCAAATATCAGGGTCACATGTTAGCTTTGGTACCTGTTAGATCGAGTTGT
    |||||
Db 421 CCGTATCCAAATATCAGGGTCACATGTTAGCTTTGGTACCTGTTAGATCGAGTTGT
    |||||
QY 421 CCGTATCCAAATATCAGGGTCACATGTTAGCTTTGGTACCTGTTAGATCGAGTTGT
    |||||
Db 421 CCGTATCCAAATATCAGGGTCACATGTTAGCTTTGGTACCTGTTAGATCGAGTTGT
    |||||
QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAATGATATTCAGTATTCAT
    |||||
Db 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAATGATATTCAGTATTCAT
    |||||
QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAATGATATTCAGTATTCAT
    |||||
Db 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAATGATATTCAGTATTCAT
    |||||
QY 541 ATTGACGACCAACAGTAGTAAATCAAGTTTATGTAATGCTATTTTATAGTGAAGATCG
    |||||
Db 541 ATTGACGACCAACAGTAGTAAATCAAGTTTATGTAATGCTATTTTATAGTGAAGATCG
    |||||
QY 541 ATTGACGACCAACAGTAGTAAATCAAGTTTATGTAATGCTATTTTATAGTGAAGATCG
    |||||
Db 541 ATTGACGACCAACAGTAGTAAATCAAGTTTATGTAATGCTATTTTATAGTGAAGATCG
    |||||
QY 601 GTTATCGTTTCTTAATGTTGTCAGGACACGAGACGCTATTTGATACAGATATGAT
    |||||
Db 601 GTTATCGTTTCTTAATGTTGTCAGGACACGAGACGCTATTTGATACAGATATGAT
    |||||
QY 601 GTTATCGTTTCTTAATGTTGTCAGGACACGAGACGCTATTTGATACAGATATGAT
    |||||
Db 601 GTTATCGTTTCTTAATGTTGTCAGGACACGAGACGCTATTTGATACAGATATGAT
    |||||
QY 661 GATGACACAAGATTATGTTTATGATCGATCTGCTGATGCTGATGCTGATGCTGATG
    |||||
Db 661 GATGACACAAGATTATGTTTATGATCGATCTGCTGATGCTGATGCTGATGCTGATG
    |||||
QY 661 GATGACACAAGATTATGTTTATGATCGATCTGCTGATGCTGATGCTGATGCTGATG
    |||||
Db 661 GATGACACAAGATTATGTTTATGATCGATCTGCTGATGCTGATGCTGATGCTGATG
    |||||
QY 721 GAATCAACTGAGAAATATTTCACTATTAAGAGCTTTAAAGCATTTGAACGTTCAATGTT
    |||||
Db 721 GAATCAACTGAGAAATATTTCACTATTAAGAGCTTTAAAGCATTTGAACGTTCAATGTT
    |||||
QY 721 GAATCAACTGAGAAATATTTCACTATTAAGAGCTTTAAAGCATTTGAACGTTCAATGTT
    |||||
Db 721 GAATCAACTGAGAAATATTTCACTATTAAGAGCTTTAAAGCATTTGAACGTTCAATGTT
    |||||
QY 781 GTTTTAGTGCTTATTTGATGACAGACAGCATTTATTTGAACAAGATTAACGTTGTCAGGA
    |||||
Db 781 GTTTTAGTGCTTATTTGATGACAGACAGCATTTATTTGAACAAGATTAACGTTGTCAGGA
    |||||

```

```

Db 781 GTTTTAGTGCTATATGATGACAGACAGCATTTATTTGAACAAGATTAACGTTGTCAGGA
    |||||
QY 841 TATGCATGACCAACAGTAAAGACGTCGATTTGCTGATTAATTAATGGATCTGTGGA
    |||||
Db 841 TATGCATGACCAACAGTAAAGACGTCGATTTGCTGATTAATTAATGGATCTGTGGA
    |||||
QY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTTCAATTTTA
    |||||
Db 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTTCAATTTTA
    |||||
QY 961 GATTATGACCAAAATTTGCTTTGTTGCTGCTAAAGAACGACAAAGATTACGTAATTTTC
    |||||
Db 961 GATTATGACCAAAATTTGCTTTGTTGCTGCTAAAGAACGACAAAGATTACGTAATTTTC
    |||||
QY 1021 CCTTACATTAATGAAGCAAGTGAACCAATTTAAACCAATTTTGAAGTCAACTTTAAT
    |||||
Db 1021 CCTTACATTAATGAAGCAAGTGAACCAATTTAAACCAATTTTGAAGTCAACTTTAAT
    |||||
QY 1081 GAAGTCGTTACTGATGCAATTTTCATGAAACCCCTACACCAACAAAGTAGAGCTTTG
    |||||
Db 1081 GAAGTCGTTACTGATGCAATTTTCATGAAACCCCTACACCAACAAAGTAGAGCTTTG
    |||||
QY 1141 AATGCTTTTATGCAACACAGTGTCTATAGAACACGACATTTGTTGTTATTTTAAT
    |||||
Db 1141 AATGCTTTTATGCAACACAGTGTCTATAGAACACGACATTTGTTGTTATTTTAAT
    |||||
QY 1201 GATGTAGAAATTAATGATTTTCTTATTAACGCTATTTGAAATCAATTCGTCGCGCT
    |||||
Db 1201 GATGTAGAAATTAATGATTTTCTTATTAACGCTATTTGAAATCAATTCGTCGCGCT
    |||||
QY 1261 TTTGGTTTGAAGGTACACCAATTCATATTATAGTGTGAAAGAGA 1305
    |||||
Db 1261 TTTGGTTTGAAGGTACACCAATTCATATTATAGTGTGAAAGAGA 1305
    |||||
RESULT 4
AP004827/c 333750 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
DEFINITION strain:MW2, section 6/10.
ACCESSION AP004827 BA000033
VERSION AP004827.1 GI:21204509
KEYWORDS Staphylococcus aureus subsp. aureus MW2
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
Yamamoto,K. and Hiramatsu,K.
TITLE Genome and virulence determinants of high virulence
community-acquired MRSA
JOURNAL Lancet 359 (9320), 1819-1827 (2002)
MEDLINE 22040717
PUBMED 12044378
REFERENCE 2 (bases 1 to 333750)
AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
Kikuchi,H.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center, 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bioente.go.jp, URL:http://www.bio.nite.go.jp/,
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES
  source
    1..333750
    /organism="Staphylococcus aureus subsp. aureus MW2"
    /mol_type="genomic DNA"
    /strain="MW2"
    /db_xref="taxon:196620"
    complement(278..619)
  gene

```


gene
CDS
/gene="WM1341"
/complement(278..619)
/gene="WM1341"
/note="ORFID:WM1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MTNTKEFDNLBOEIOVEAKNKTLSGESEAKPYLDEHYSKIID
EKQVNDITGNIDFDKLNQYPPVPMNFKERYDMIRKTHFMGYROMKTFKELLNMA
SYTRLNKNQV"
complement(624..1283)
/gene="hth"
complement(624..1283)
/gene="hth"
/note="ORFID:WM1342"
/codon_start=1
/transl_table=11
/product="endonuclease-like protein"
/protein_id="BAB95207.1"
/db_xref="GI:21204511"
/translation="WVSKKALEMIDVIANMPPDAECLEKHDNPELTITAVLSAQC
TVLVARVTELEFKTKTPEDYLAVSDEELMNDIRSIGLRNKAIRKLCOSLIDQFN
GEIPQTHKELESAGVGRTANVMSVAFDEPSLAVDTHVERVSRLGINRMKDNVRQ
VEDRLCSVIPRDMRMRSHQILFEGRYHCLARKPKCDICPLEDCREGOKRYKASLKE
A"
complement(1273..1959)
/gene="WM1343"
complement(1273..1959)
/gene="WM1343"
/note="ORFID:WM1343
hypothetical protein, similar to chromosome replication
initiation protein dnaD"
/codon_start=1
/transl_table=11
/protein_id="BAB95208.1"
/db_xref="GI:21204512"
/translation="MDXYOLKARPVIRRELLDHYSDIGLDEODLVILLKIYASETS
NKPSIELLOKSTWQPRDITMVIONLQRELELOVOKDEGRTEVNNIDPFEXL
SHLAKQSMETREONSKSEFKQLFVLEDTFARPLSPLELELNQWIDVCKHDSAIIO
AALDEANLNLSEFKYMDRILLNMKKNVKTIDSRKIREKFNKPKMTYKTVKFD
WLNGENLDGK"
complement(2287..3579)
/gene="asns"
complement(2287..3579)
/gene="asns"
/note="ORFID:WM1344"
/codon_start=1
/transl_table=11
/product="asparaginyl-tRNA synthetase"
/protein_id="BAB95209.1"
/db_xref="GI:21204513"
/translation="WKTTIKQAKDHLNODVTIGAWLTKNRSSGKIAPLQLRDGTFMQ
GVVXKSEVDEVEFKLAKETIOESSLYVGTITEDNRSDLGEMQKSEIVSEADPY
ITPKNHGTEFLMHLRLRSKQCHAVMKIRNEVIRATYEFENKDGFTKVDPILTAS
APEGSELFHTKTFPODAFLQSGOLYLEAAMAHGKVPFSGPTRAKRSKRRLLIE
FMMEGEMAFTHNAESLEFIOQOYTHVYKSVLENKELKIERPISLKEKVAITPFR
ISYDDAIFLKSSEGFDDLEMGEDFGAPHEIATANHIDLPVFTNTPKIKPPYQPNP
ENETVLCADLIAPGYGELIGGSEKRVDELLEQDRAVEHGHDEAVSYVDLRRYGS
VPHCGFGLERTYAMISGVEHVRETAFPRILNRLYP"
complement(3901..6594)
/gene="ding"
complement(3901..6594)
/gene="ding"
/note="ORFID:WM1345"
/codon_start=1
/transl_table=11
/product="probable ATP-dependent DNA helicase ding"
/protein_id="BAB95210.1"
/db_xref="GI:21204514"
/translation="MGATYAVVDDLETTGNQLDPDDIIOIGITFVRNNOIIDTYHSMI
RTNLEIPPIQALTSIEENMLQOAPYFNQVAOETDYKIKDCIFVAHNVDFLNFIKKA

gene
CDS
/gene="WM1346"
complement(6618..7589)
/gene="WM1346"
/note="ORFID:WM1346"
/codon_start=1
/transl_table=11
/product="truncated hypothetical protein"
/protein_id="BAB95211.1"
/db_xref="GI:21204515"
/translation="MSKYSODVLOQLYKKNKPNYISGQSIASBSLNSRTAVKAVDQK
LEGCKIDSVNHHGHLLOLPDIWOGIIDQYKSSALPDFSEKVDSDISTOLAARSL
VGNQSSFILSDQTKGGRFRNRHSSKCGQUMSVYLRPVASMSKRLPITALG
IKDAIQHSDQEKVMPNDIYIDNGKCGFLEEVANNDDGEALICGIGLITQOLE
NFDESIRHRATSVQLHDKNKLDYQFLQELQELKRVNQLPFPSELRREYIAASN
IWNRTILFTEENKQFKGOAIDLDYDGYLVLRDEAGESHRLISADIDF"
complement(7576..8778)
/gene="WM1347"
complement(7576..8778)
/gene="WM1347"
/note="ORFID:WM1347
hypothetical protein, similar to poly(A) polymerase"
/codon_start=1
/transl_table=11
/protein_id="BAB95212.1"
/db_xref="GI:21204516"
/translation="MDKSLFEQARPILEQIDNGFEAVYVGGSVRDYVWGRINIDIDI
TTSATPDEIESIFSHTIPGKEHGHTINNVFNDENEVTTFRAEEDYVDHRRSGVTFV
RDLYDELQORDFTNANAAMDYAKLYDYDFGQDINNRIITVGAIEEFQEDALRMS
RLRFOSQSFQIATEFEFAMRIQADIKFLSIBRIYELTRLMGINKVSEKFNILKS
LKAFNMPYFELDMNOJINPEAIDELLIAVSVPDINYSIKPLKXISNRQVNDIO
YIQINNALPSITTKFQLKRFYVDYTHLIRKNVAAADVIKANDIDGHEPLIYNLQTTD
ETLHRLPMHNRKDMVNGVLAHLNAKSGPWLKDVLLQIETAIYTGKAVNEETELK
WVDNHVKI"
complement(8783..9925)
/gene="WM1348"
complement(8783..9925)
/gene="WM1348"
/note="ORFID:WM1348
hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"
/translation="MKIGITCPYSGMGSGQIATELGIKLAERGHVHTTSNIPRIIR
KPLPMIFHOVENOYAVFOYPDYITLSTKIAEYIKESGLDLHMHYAPPAICGIL
AREMSGDKIKMTLTGHDITVIGYDLSLQAKIKGIEKSDIVTSVSKSLAODEHEIT
ETNKEITIPYNVRENREPTKNTALKSQFGAIREPKVLIHNSRQVRRIDETIETP
AKYREKIPSKLILLGDGPELVPKROLIRELNEEDYVLEKQDCYSEFYQULSDVLL
SEKESFGLLEAMKTVGVPISGNAGGIREVIRKHEGTGVVDVGGCDSNDAIRLLE
DKVLVNLQKNMLADIAERFSESLITTDQYEVYQQMLNEHNSKGE"
complement(10167..10484)
/gene="WM1349"
complement(10167..10484)
/gene="WM1349"
/note="ORFID:WM1349"
/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMQREVDVYIGQFTGYFSPLANLARLEVEGELAREIN
HTYGEKKKKDSEBANTIKAEIDGNLFVLLCLANSMGIDMTESNEMEMFNTRDKRRF
EKK"
complement(10820..11518)
/gene="MM1350"
complement(10820..11518)
/gene="MM1350"
/note="ORFID:MM1350"
/cdd_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIFPMIITYEVILWPLMAQHKVSNKEYSOVRSSTGKTR
EVALLEIHANGIYDVEVVKGEFLTDHYOPKKVYLSLSPANDRPSVATLAIHAEVG
HAIQDHQGWELFRNALVPVANLSSLSLMIILITLILALISAFGSTALMIGAGLM

Query Match 98.9%; Score 1290.6; DB 1: Length 333750;
Best Local Similarity 99.3%; Pred. No. 2,3e-214;
Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTATGCTATGAGTGGCCCTAATGTAGCTAATCTACAAATTTT 60
DB 24351 ATGACTAAACCTATAGTATGCTATGAGTGGCCCTAATGTAGCTAATCTACAAATTTT 24292
QY 61 AATAGATAGTGGAGAACGCTGTTGATGTTGAAGATACGCCAGGTGAACACGAGAT 120
DB 24291 AATAGATAGTGGAGAACGCTGTTGATGTTGAAGATACGCCAGGTGAACACGAGAT 24232
QY 121 CGTATTTATCTTCAGGTGATGTTTACACATGATTTCAATATTTATGATACAGTGT 180
DB 24231 CGTATTTATCTTCAGGTGATGTTTACACATGATTTCAATATTTATGATACAGTGT 24172
QY 181 ATTGAATGTTGATGATCACCATTCCAAACCAAAATTAGACGCGACAGAAATCCGCAT 240
DB 24171 ATTGAATGTTGATGATCACCATTCCAAACCAAAATTAGACGCGACAGAAATCCGCAT 24112
QY 241 GATGAACGCGATGTTATTTATGTTTAACTGCTGATGAGATTTGACACAAAGCAT 300
DB 24111 GATGAACGCGATGTTATTTATGTTTAACTGCTGATGAGATTTGACACAAAGCAT 24052
QY 301 GAAATGCTGCTCAAAATTTTATCAAAATCTAAAAACCGGCTGATTAAGCGTTAA 360
DB 24051 GAAATGCTGCTCAAAATTTTATCAAAATCTAAAAACCGGCTGATTAAGCGTTAA 23992
QY 361 GTAGATTAATGGAATGCGGTACAGACGCTATGATTTCTATTCATTAGATTTGGTGA 420
DB 23991 GTAGATTAATGGAATGCGGTACAGACGCTATGATTTCTATTCATTAGATTTGGTGA 23932
QY 421 CCCTATCCAAATACAGGTCACATGTTAGTCTTGAGTCACTTGTAGATGAGTGT 480
DB 23931 CCCTATCCAAATACAGGTCACATGTTAGTCTTGAGTCACTTGTAGATGAGTGT 23872
QY 481 TCTCATTTTGGTGAAGAGAAGATCCTTATGATGAAGATCAATTCGACTATCCATT 540
DB 23871 TCTCATTTTGGTGAAGAGAAGATCCTTATGATGAAGATCAATTCGACTATCCATT 23812
QY 541 ATTGGACGACCAAGCTAGTAAATCAAGTTTGTAAATCCCTTTTATAGTGAAGATCC 600
DB 23811 ATTGGACGACCAAGCTAGTAAATCAAGTTTGTAAATCCCTTTTATAGTGAAGATCC 23752
QY 601 GTTATGTTTCTAATGTTGACAGGACACAGAGAGCTATTGATACAGAGTATTAGTAT 660
DB 23751 GTTATGTTTCTAATGTTGACAGGACACAGAGAGCTATTGATACAGAGTATTAGTAT 23692
QY 661 GATGACAAGATTATGTTTAAATCGATCTGCTGATGCGTAAAAAAGAAAAAGTATAT 720
DB 23691 GATGACAAGATTATGTTTAAATCGATCTGCTGATGCGTAAAAAAGAAAAAGTATAT 23632
QY 721 GAATCAACGAGAATAATTCAGATTAAAGCTTTTAAAGCGATTGAACGTTCAAAATGTT 780

DB 23631 GAATCAACGAGAATAATTCAGATTAAAGCTTTTAAAGCGATTGAACGTTCAAAATGTT 23572
QY 781 GTTATGTTGTTTATGATGACAGAACAGCATATTGAACAAATTAAGCTTTGCAGCA 840
DB 23571 GTTATGTTGTTTATGATGACAGAACAGCATATTGAACAAATTAAGCTTTGCAGCA 23512
QY 841 TATGACATGACAAAGGTAAAGCAGTGTGATTTGCTGAATTAATGATGATCTGTGAA 900
DB 23511 TATGACATGACAAAGGTAAAGCAGTGTGATTTGCTGAATTAATGATGATCTGTGAA 23452
QY 901 AAGATAGTAAACGATGGAAGAAATTTGAAGATGAGTACGTTAAAGATTCATTTTAA 960
DB 23451 AAGATAGTAAACGATGGAAGAAATTTGAAGATGAGTACGTTAAAGATTCATTTTAA 23392
QY 961 GATTATGACAAATTCCTTTTGTGCTGCTGAAGAACGACCAAGATTACGATATTTC 1020
DB 23391 GATTATGACAAATTCCTTTTGTGCTGCTGAAGAACGACCAAGATTACGATATTTC 23332
QY 1021 CCTTACATTAATGACCAAGTGAAGAAACCATTAAGAAAGTTCACATTTAAAT 1080
DB 23331 CCTTACATTAATGACCAAGTGAAGAAACCATTAAGAAAGTTCACATTTAAAT 23272
QY 1081 GAAGTGTGCTGATGCAATTTCCATGAAACCCCTACACCAACACAAAGGTACAGTTTG 1140
DB 23271 GAAGTGTGCTGATGCAATTTCCATGAAACCCCTACACCAACACAAAGGTACAGTTTG 23212
QY 1141 AATGCTTTTATGACACACAAAGTGTGATGAAACCCGACCATTTGTTGATTTGTTAAT 1200
DB 23211 AATGCTTTTATGACACACAAAGTGTGATGAAACCCGACCATTTGTTGATTTGTTAAT 23152
QY 1201 GATGTAAGATTAATGCAATTTTCTTAAACGCTATTAGAGATCAATCCGTCGCCCT 1260
DB 23151 GATGTAAGATTAATGCAATTTTCTTAAACGCTATTAGAGATCAATCCGTCGCCCT 23092
QY 1261 TTGTTGTTGAAGTACACCAATTCATATTATGCTGGAAGAGA 1305
DB 23091 TTGTTGTTGAAGTACACCAATTCATATTATGCTGGAAGAGA 23047

RESULT 5
AF270032/c 3269 bp DNA linear BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1042f07 genomic
DEFINITION
ACCESSION AF270032
VERSION
KEYWORDS
SOURCE Staphylococcus epidermidis
ORGANISM Staphylococcus epidermidis
REFERENCE 1 (bases 1 to 3269)
AUTHORS Kimmerly,W.J., Taylor,J.David., Neelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3269)
AUTHORS Taylor,J.David., Kimmerly,W.J., Neelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1..3269
/organism="Staphylococcus epidermidis"

/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1042f07"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN

Query Match 70.5%; Score 919.4; DB 1; Length 3269;
Best Local Similarity 81.5%; Pred. No. 1.3e-149;
Matches 1064; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTACTATTGTAGTAGGCGCTAAATGTAGTAATCTACAATTTT 60
DB 2804 ATGACTAAACCGATAGTACTATTGTAGGAAACCAAAATGTAGGAAATCTACAATTTT 2745
QY 61 AATAAATAGTTGGGAACGCTTTCGATTTGTTGAAGATACCCGAGGTAAACAGAGAT 120
DB 2744 AATAAGTTGTGGGGAACGATATCTATCGTAGAAGATACGCCAGGCGCTTCTCGTGAT 2685
QY 121 CGTATTTATTCCTTCAGGTGAGTGGTTAACACATGATTTCAATATTATTGATACAGGTGT 180
DB 2684 CGTATTTATTCATCTCGGCGAGTGGTTAACTCATGATTTTAACATCATTCATACAGGTGT 2625
QY 181 ATTGAATTTGGTATGCACACATTTCCAAACACAATTTAGAGCGCAGGACAGAAATCCGCATA 240
DB 2624 ATTGAATTTGGAGATGCTCTCTTTCAAACTCAATCCGTCAGACAGGACAGAAATGCAATTT 2565
QY 241 GATGAAGCGGATGTATATTTTATTTTATGTTAACGTGCGCGTAAGGATGTGACAAAGCGAT 300
DB 2564 GAGAAGACAGATGTATCATCTTTTATGTCATTTATGGAAGGACTTACACAAAGTGAC 2505
QY 301 GAATGTCGCTCAATTTTATACAAATCTAAAAACCGGCTGTATTAGCGTTAAACAA 360
DB 2504 GAATGTCGCGCAAAATGCTTTATATATCTAAGAAACCGTGTATTAGCTGTGAATTA 2445
QY 361 GTAGTAATATGGAATGCGTACAGACGTGTATGATTTCTATTTCAATTTAGATTTGGTGAA 420
DB 2444 GTTGATTAATCTGAATGCGTATGATATCTATGATTTCTATTTAGCTTTGGAGAT 2385
QY 421 CCGTATCCAAATTCAGGCGTACATGTTAGTCTTTGGGACTTTGTAATACAGTGT 480
DB 2384 CCATATCCATTTTCTGTTTCACATGATTAAGCTTTGGAGATTTCTATACAGTGT 2325
QY 481 TCTCATTTTGGGAGAGAGAAAGATATCTTATGATGAATGATACATTCGATTCATC 540
DB 2324 GAAACCTTTAATAAAGAAATCAGAAATCTTTATGACGAAGATACGATCTTTCTATAC 2265
QY 541 ATTGAGCAGCAAAACGTAGGTAAATCAAGTTAGTAATGCTATTTTAGTAGAAGATCCG 600
DB 2264 ATCGGTAGACCTAATGTTGGTAAATCTAGCTTGGTCAATGCTATTTTAGCGAAGACGT 2205
QY 601 GTTATCGTTTCTAATGTTGCGAGGACACGACAGACGCTATTGATACAGATATGAT 660
DB 2204 GTTATGTTCTAATGTTGCTGCTGACAACTCGAGATGCCATTCATCCGATGACTCTTAT 2145
QY 661 GATGACAGAGATTTATGTTTATTCGATCTGCTGATGCTGTAAGGAAAGAAAGATATAT 720
DB 2144 GATGACAGAGATTTATGTTTATGTTGATCTGCTGATGAGAAAGAAAGGTTAGGTAT 2085
QY 721 GAATCACTGAGAAATATTCACTATTAAAGCTTTAAAGCGATTGAACGCTCAATGTT 780
DB 2084 GAATGACTGAGAAATATTCTGATTAGCTGATTAAGGATGAGGCTTCAAGATTA 2025
QY 781 GTTTTAGTGTATTGATGACAGACAGGCAATTTGAACAAGATTAACGTTGTGACAGA 840
DB 2024 GTATTAGTGTATTGATGACAGACAGGATTAATTAAGCAAGATTAACGTTAGTGTGCG 1965
QY 841 TATGACATGACAAAGGTAAAGCGTGTGATTTGCTGTAATTAATGGAATCTGCTGGA 900
DB 1964 TATGACATGACAAAGGTAAAGCGTGTGATTTGCTGTAATTAATGGAATCTGCTGGA 1905
QY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTTCCAAATTTTA 960
IIIIIIIIII II IIIII IIII II II IIIII IIIIIIIIIII

DB 1904 AAAGATAGTAAACATGAAAAAATTCAGTGTGATGTAGAAATGAAATTTCAATTTT 1845
QY 961 GATTATGCAAAATTTGCTTTGTCGTCTAAAGAACGCAAGATATTACGATTAATTC 1020
DB 1844 GATTATGCTCAAAATGCGCTTCTATACAGCAAAAGAGGCTTAAGATTTAAAAACATTATTC 1785
QY 1021 CCTTACATTAATGAAGCAAGTAAAAACCTAATAAAACGTTTCAAAAGTTCAACTTTAAT 1080
DB 1784 CCTTATATCAATCAAGCCGAGTGAATAATCTAATAAAGCGTGCACAAAGTTCTACAGTAAT 1725
QY 1081 GAATGCTGTACTGATGCAATTTTCCATGAACCTTACACACAGACAGCAAGGTAAGCTTTG 1140
DB 1724 GAATGCTGTACTGATGCAATTTTCCATGAATCCAAACCTAATCAGCAAAAGTGAAGACTT 1665
QY 1141 AATGCTTTATGCAACCAAGTGTGTATAGAACACCGACATTTGTTGATTTGTTAT 1200
DB 1664 AATGATTTCTATACCACTGAGTGTGAATGAACCAACGACATTTGATATTGTCAT 1605
QY 1201 GATGAGAAATTAATGCAATTTTCTTATAAACGCTATTATTAGAAATCAATCCGTGCGCT 1260
DB 1604 GATGTTGAATTAATGATTTTCTTATAGAGATATTAGAAATCAATTAAGTATGCT 1545
QY 1261 TTTGCTTTGAAGTACACCAATTCATATTATAGCTCGAAGAGA 1305
DB 1544 TTTGCTTTGAAGAACACCTATTCTATTTATTTCCAAAGAAAAAGA 1500
RESULT 6
AX145350/c 3269 bp DNA linear PAT 31-MAY-2001
LOCUS AX145350
DEFINITION Sequence 4072 from Patent WO0134809.
ACCESSION AX145350
VERSION AX145350.1 GI:14283915
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Kimmerly, W. J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4072 17-MAY-2001;
FEATURES
source 1. 3269
Location/Qualifiers
1. 3269
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN
Query Match 70.5%; Score 919.4; DB 6; Length 3269;
Best Local Similarity 81.5%; Pred. No. 1.3e-149;
Matches 1064; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATAGTACTATTGTAGTAGGCGCTAAATGTAGTAATCTACAATTTT 60
DB 2804 ATGACTAAACCGATAGTACTATTGTAGGAAACCAAAATGTAGGAAATCTACAATTTT 2745
QY 61 AATAAATAGTTGGGAACGCTTTCGATTTGTTGAAGATACCCGAGGTAAACAGAGAT 120
DB 2744 AATAAGTTGTGGGGAACGATATCTATCGTAGAAGATACGCCAGGCGCTTCTCGTGAT 2685
QY 121 CGTATTTATTCCTTCAGGTGAGTGGTTAACACATGATTTCAATATTATTGATACAGGTGT 180
DB 2684 CGTATTTATTCATCTCGGCGAGTGGTTAACTCATGATTTTAACATTCATTAACAGGTGT 2625
QY 181 ATTGAATTTGGGATGACCATTCACCAACACAATTAAGAGCGCAGACAGAAATCCGCATA 240
DB 2624 ATTGAATTTGGAGATGCTCTCTTTCAAACTCAATCCGTCAGACAGGACAGAAATGCAATTT 2565
QY 241 GATGAAGCGGATGTATTTTATTTTATGTTAACGTGCGTAAGGATTTGACAAAGCGAT 300
IIIIIIIIII II IIIII IIII II II IIIII IIIIIIIIIII

D 2564 GAAGAACAGATGTCATCATTTTATGCTAAATGTTAGAGAGAGACTTACACAAAGTGAC 2505
Q 301 GAAATGGTCGCTCAAAATTTTATACAAATCTAAAAAACCGGTCGATTAGCGTTAACAAA 360
D 2504 GAAATGGTCGCAAAATGCTTTTAAATCTAAGAAACCTGTTGATTAGCTGGAATAAA 2445
Q 361 GTAGATTAATGGAATGCGTACAGACGCTATGATTTCTATTCATTAGCATTTGGTGAA 420
D 2444 GTTGATTAATCTAAAGCGCTAAATGATATGATTTCTATTCATTAGCTTTGGAGAT 2385
Q 421 CCGTATCCATATCAGGTCACATGCTTGTAGCTTGTGACTTGTGATGAGATGAGTGT 480
D 2384 CCAATTCCTATTTCTGTTCCATGCAATGAGATTGAGACTTGGAGATTTGCTAGATTCAGTTGTT 2325
Q 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAGATACAAATTCAGTATCCATT 540
D 2324 GAAACCTTTTAAAGAAATCAGAAAGATCCCTTATGACAGATATGAGATACGCTTTTCTATC 2265
Q 541 ATTGAGACGCCAAACGATTAATCAAGTTTATGTAATCTATTTAGTGAAGATCCG 600
D 2264 ATGCGTAGACCTAATGTTGGTAAATCTAGCTTGGTCAATGCTATTTTGGCGAAGACGT 2205
Q 601 GTTATCGTTTCTAATGTTGACGAGACGAGAGACGCTTATGATACAGATATAGTTAT 660
D 2204 GTTATGTGCTCTAATGTTGCTGCTGACACCTGAGATGCCATTGATACCGAGTACTTTAT 2145
Q 661 GATGACAGATTAATGTTTAAATCGATCTGCTGTAATGCGTAAAAAGGAAAGATAT 720
D 2144 GATGACAGATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2085
Q 721 GATCAACGAGAAATTAATTAAGATTAAGGCTTTAAAGCGATTGACGTTCAATGTT 780
D 2084 GAATCAACGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2025
Q 781 GTTATGCTGTTTATGATGAGAGAAAGGATTAATGACATTAATGACATTAATGACAT 840
D 2024 GTTATGCTGTTTATGATGAGAGAAAGGATTAATGACATTAATGACATTAATGACAT 1965
Q 841 TATGACATGACAAAGGTAAGGATGCTGATGCTGTAATTAATGATGATGATGATGAT 900
D 1964 TATGACATGACAAAGGTAAGGATGCTGATGCTGTAATTAATGATGATGATGATGAT 1905
Q 901 AAGATAGTAAAGGTAAGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 960
D 1904 AAGATAGTAAAGGTAAGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1845
Q 961 GATTATGACAAATGCTTTTGTGCTGTAAGAGACGACAGATTAATGATGATGATGAT 1020
D 1844 GATTATGCTCAATTCGCTGCTGATCAGCAAAAGAGGCTAAGATTAATTAATTAATTAAT 1785
Q 1021 CCTTACATTAATGACAGAGTGAACCATTAATTAATTAATTAATTAATTAATTAAT 1080
D 1784 CCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1725
Q 1081 GAATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
D 1724 GAATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
Q 1141 AATGCTTTTATGACAGAGTGAACCATTAATTAATTAATTAATTAATTAATTAAT 1200
D 1664 AATGATTTCTATCAACTCAGGTTGCAATTAATTAATTAATTAATTAATTAATTAAT 1605
Q 1201 GATGATTAATTAATGATTTTCTTAAACGCTATTAATTAATTAATTAATTAATTAAT 1260
D 1604 GATGATTAATTAATGATTTTCTTAAACGCTATTAATTAATTAATTAATTAATTAAT 1545
Q 1261 TTTGCTTTGAGAGTACCAATTCATTAATTAATTAATTAATTAATTAATTAATTAAT 1305
D 1544 TTTGCTTTGAGAGACCACTATTCATTAATTAATTAATTAATTAATTAATTAATTAAT 1500

RESULT 7
AB016747/c

LOCUS AB016747 300892 bp DNA linear BCT 01-JAN-2003
DEFINITION Staphylococcus epidermidis ATCC 12228, section 4 of 9 of the complete genome.
ACCESSION AB016747 AB015929
VERSION AB016747.1 GI:27315369
KEYWORDS Staphylococcus epidermidis ATCC 12228
SOURCE Staphylococcus epidermidis ATCC 12228
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 300892)
Zhang, Y., Ren, S., Li, H., Fu, G., Lu, L., Lu, G., Jia, J., Tu, Y., Qin, Z., Chen, Z. and Wen, Y.
Direct Submission
Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China
location/Qualifiers
1. 300892
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
113. 346
/locus_tag="SE0907"
113. 346
/locus_tag="SE0907"
/codon_start=1
/transl_table=11
/product="acyl carrier protein"
/protein_id="AA004504.1"
/db_xref="GI:27315370"
/translation="MENFDKVDIIVDRIGVDADKVTEDASFKDGLGADLIDAEIWM
ELEDFEGTEIPDEAEKINIVGDAVYINSLEK"
461. 1198
/locus_tag="SE0908"
461. 1198
/locus_tag="SE0908"
/transl_table=11
/codon_start=1
/product="Rnase III"
/protein_id="AA004505.1"
/db_xref="GI:27315371"
/translation="MANOKKKEVNHPOOKFTDPMKMSLGHFKNKIDLYOAPSHSFI
NDNMNRLEHNERLEFLGDVAVLEIVSRILEDPHLPBENLTKMRATTVCEPSLIF
ANKIKNELNILLGKEGKEKGTGRPSLISDAFEAGALYLDGDLFSVTFAPKVIPE
YVEDDELIVGVDFKTOFOEYVHSONKGDVYQLIKEGPAHRLTFSEVILENKAAVE
GKRTKTESPOKABQAYKLMKKSL"
1325. 4894
/locus_tag="SE0909"
1325. 4894
/locus_tag="SE0909"
/codon_start=1
/transl_table=11
/product="chromosome segregation SMC protein"
/protein_id="AA004506.1"
/db_xref="GI:27315372"
/translation="AWYLSIDAIGFKSPADHTNVOFDKGVTAIVPNSGKSNITDA
IKWILEOSAKSLRGAKMEDIIIFSGAEHKKAOFAVULKLDHNSQKLOIDSEELVT
RLYRSGESEYVYNNDRARLKDIIDFLSGLCKEKFSLISGRVDEILNARINRRO
IIFESAGVLYKKRKRAESTOKLDHPEFDNLNRYVDILYDLEGREPKAASATAKEKO
LSKEMQSDVITVSDIDHTEDNOLDEHLNLKSOQAEKESQAOIOLQLOKIKGK
ROONDYDIEKLNELVKAATENYEQLSGKLVLEERKNOSETNARYEEELNDLEQID
SINKERAOEKLAELEKNOKOLNEVOELESLLYISDEOHKELEIKNSYTTLMSE
OSVDNNDIREFLEHTINENPAKSKRLSDRLVEAPNOLQIDQONTLOKEYOSSKME
KVRONILOOLEOLDSKRLISEYENKLYOAYRNEKLSRISLSLATQEPEDYTFENYV
KHILKAKDELRGIRHVAEYIVPSEMTQALITFAGASLQHYIVNENDRQAIOLY
KQGLCRATFLPLNVLOPRHVADINDVARGSGFTINISDAIWSAKONTINELG
NTIIVENLKHANELARIRYRPIVTLGDDVNPVPGSMGSGARXKSLISQKDELST
MRNOLDEYOROTAEFEROPREKQTOAEQSEGYFSASOOYNNILKEOVHHLELDLKL
TOEHLNKEHEEPENKNGYOSDKSKETLKEKONMLIELOOLKEDEIDERYTOS
KEKASTHOTLOOLHOKQSLDAVYKERSOKOYVREKLOKLOSDSROKIEVEXIKL
FNSDEMCKDAPEKLEQIOTOCENYQNNLOLSEIKOQRKDLNEXIETINESLOKCH
ODLSTENHYODIKAKOSKLDVLIINAHIDLNTYOLTVERRAMEYDSQETIDNLKPK

Db	297004	TTTGAAATTTGGAGATGCTCTTTTCAAACTCAAAATCCGTGCACAGGACAGAAATACGAAT	296945
Qy	241	GATGACAGCGGATGTTATTTATTTTATGTGTTAACTGCGGTGAAGGATTGACACAAACCGAT	300
Db	296944	GAAGACAGATGTCATCATTTTTTTATGTGCATCTTAGAGAGAGACTTACACAAAGTGAC	296885
Qy	301	GAAATGGCGCGTCAAAATTTTATACAAATCTAAAAACCGGTGCTATTAGCGTTAACAA	360
Db	296884	GAAATGGTCGCACAAATCTTTTATTAATCTTAGAAACCTGTTGTAATTACTGTGAAATTA	296825
Qy	361	TGATATTAATATGAAATATCGTACAGACGCTGTATGATTTCTATTCAATTTAGATTTGTGAA	420
Db	296824	GTTGATTAATCTTGAATATCGTAATGATATATCATGATTTCTATTCTTTAGCGTTGAGAT	296765
Qy	421	CCGTAATCCAAATATCAGGGTCACATGCTTTAGGCTCTTGGTACTTGTATGATCAGTTGT	480
Db	296764	CCATATCTTAATTTCTGTTTCCATGATGATTAGACTTGGAGATTTGCTATCATCAGTTGCT	296705
Qy	481	TCTCATTTTGGGAAGAGAGAAATATCCTTTATGTATGAAAGATACATTCGACTATTCAT	540
Db	296704	GAAATCTTTAATTAAGAAATCAGAAATATCCTTATGACGAAGATACGATACGTTCTTATAC	296645
Qy	541	ATTGACAGCCAAACGTAAGTAAATCAAGTTTATGTAATATGCTATTTTATAGTGAACATCC	600
Db	296644	ATCGGTAGACCTAATGTTGTAATCTAGCTTGGTCAATGCTATTTTATAGCGAAGACGT	296585
Qy	601	GTTATCGTTTCTAATGTTCAGGCAACAGAGACGCTATTGATACAGATATAGTAT	660
Db	296584	GTTATTTGTCTAAATGTCTGGTCAACTCGAGATGCGCATTTGATACCGATACCTTTAT	296525
Qy	661	GATGACACAAGTTATGTTTAAATGATCTGCTGTATGCCGTAATGAAAAAGAAAGATATAT	720
Db	296524	GATGACACAAGTTATGTTATGATTTGATTTATCTGCTGAATGAGAAAAAAGTAAGGTAT	296465
Qy	721	GAATCAACTGGAATAATTTAGTATTAAGAGCTTTAAAAAGCATTGAACGTTCAATCTT	780
Db	296464	GAATCGACTGAAAAATATTTCTGATTTAGGTGATTTAAAAAGCATTTGAGCGTTCAAGAGTA	296405
Qy	781	GTTTATGTTGTTATTTATGATGACAGAACAGGCATTATTGAACAAGATTAACGTTGCAGA	840
Db	296404	GTATTATGATTTATTCGATGCTGACACAGGTAATAATTGAACAAGATTAACGTTAGCTGCG	296345
Qy	841	TATGACACTGAACAAGTAAAGCATCGATTTGCTGTAATAAATGGATGCTGTGGAA	900
Db	296344	TATGACACTGAGGAAGGTAAAGCTATTTGTCATTTAGTAAATAATGGATACAGTTGAA	296285
Qy	901	AAAGTACTAAACAGATAGAGAAATTTAGATGAAGTACGTAAAGAAATTCCAATTTTAA	960
Db	296284	AAAGTATTAAGACATATGAAAAAATTCCTACTGATGATGTTAGAAATGAATTTCAATTTTAA	296225
Qy	961	GATTATGACAAATTTCTTTTGTGTCTCTAAAGAACGCACAAAGATTACGTACATTATTC	1020
Db	296224	GATTATGCTCAAAATGCGTCTGATACACAAAAGAGGCTTAAGATTTAAAAACATTATTC	296165
Qy	1021	CCTTATCATTAATGAAGCAAGTGAACCATATAAAACGTTTCAAAAGTTCAACTTTAAT	1080
Db	296164	CCTTATATCATATCAAGCCAGTGAATAATATAAAAGCGTGTCCAAAGTCTACACTAAAT	296105
Qy	1081	GAAAGCTGTACTGATCAATTTCCATGGAACCTTACCAACAAGCAAAAGGTGACGTTTG	1140
Db	296104	GAAAGTTTACTGATGCGCATCTCTATGTAATCCAAACCTTACTGACAAAGGTGAAAGACTT	296045
Qy	1141	AATGCTTTTATGCAACAAGTTGCTATAGAACCAACGACATTTTGTGTATTTGTTAAT	1200
Db	296044	AATGATTTCTATPACACATCTCAGTTGCAATTTGAACACGACATTTGTATTTTGTTCAT	295985
Qy	1201	GATGATGATTTAATGCAATTTTCTTATATAAGCGTATTTTAGAGATTCAAATCCGTGCCGT	1260
Db	295984	GATGTTGATTTAATGCAATTTTCTTATATGAGATATTTTAGAAATCAAAATAGTATATGCT	295925
Qy	1261	TTTGGTTTGAAGGTACACCAATTCATTTTATAGCTCCAAAGAGA	1305
Db	295924	TTTGGTTTGAAGGACACCTATTCATATTTATTTCCAAAGAGAGAA	295860

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AE017002	Bacillus cereus ATCC 14579 section 5 of 18 of the complete genome.	AE017002	AE016817	AE017002.1	GI:29894935	
REFERENCE	Bacillus cereus ATCC 14579					
AUTHORS	Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.					
TITLE	1 (bases 1 to 304680)					
JOURNAL	Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatal,V., Bhattacharyya,A., Resnik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fongstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.					
MEDLINE	Genome sequence of Bacillus cereus and comparative analysis with					
PUBMED	Bacillus anthracis					
REFERENCE	Nature 423 (6935), 87-91 (2003)					
AUTHORS	2 (bases 1 to 304680)					
TITLE	Candelson,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.					
JOURNAL	The number of ribosomal RNA operons in Bacillus cereus					
MEDLINE	Unpublished					
PUBMED	3 (bases 1 to 304680)					
REFERENCE	Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatal,V., Bhattacharyya,A., Resnik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fongstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.					
AUTHORS	Direct Submission					
TITLE	Submitted (12-MAR-2003) INRA, GenetiQue Microbienne, Domaine de					
JOURNAL	Vilvert, Jouy en Josas 78352, France					
MEDLINE	location/Qualifiers					
PUBMED	1..304680					
REFERENCE	/organism="Bacillus cereus ATCC 14579"					
AUTHORS	/mol_type="genomic DNA"					
TITLE	/strain="ATCC 14579"					
JOURNAL	/db_xref="ATCC:14579"					
MEDLINE	/db_xref="taxon:226900"					
PUBMED	68..427					
REFERENCE	/locus_tag="BC1239"					
AUTHORS	68..427					
TITLE	/locus_tag="BC1239"					
JOURNAL	/codon_start=1					
MEDLINE	/transl_table=1					
PUBMED	/product="hypothetical Membrane Spanning Protein"					
REFERENCE	/protein_id="AAP08224.1"					
AUTHORS	/db_xref="GI:29894936"					
TITLE	/translation="MLRKLLYLFVPLVYMLISMTKIDMALFLGVYFTLSMLNRI					
JOURNAL	QEOGILICLFESIELEFSIVLNLDELFLRHISFTLMLKMTNIIOLAICIVFVYV					
MEDLINE	KEIYGKOTVFEDNRKWE"					
PUBMED	551..2041					
REFERENCE	/locus_tag="BC1240"					
AUTHORS	551..2041					
TITLE	/locus_tag="BC1240"					
JOURNAL	/codon_start=1					
MEDLINE	/transl_table=1					
PUBMED	/product="L-lactate permease"					
REFERENCE	/protein_id="AAP08225.1"					
AUTHORS	/db_xref="GI:29894937"					
TITLE	/translation="MLALPIIMIFICLFLFKQTSLNSSLSIVSVGIVLSPFOI					
JOURNAL	LGISTVHATIGKWLICIVGVLFFGFLFLPLMKMGIDIVARPLEVYDRLLOM					
MEDLINE	LMGCIQPLIESVGFGEFVNAIPFISLFLKPOVVLISFGLASWGMATGT					
PUBMED	IGSILNMPLETLGTNTALSLIPMEAYEVLISLVHVGOMQVIEKKEGVGFLESL					
REFERENCE	GIYSNAYVTEELAGLISIVTIEFGFIIIRLKKGSONITEHAATERSVITIKI					
AUTHORS	SPYELFVLCILSLRVPALHDLFRSYAVLDKSYKYLELSPGMIGMCFTIIR					
TITLE	FRIPSNITKOSLSQTIKOWIPFAITTTMTFAISLGMGSHSLAKTAGDTCTFPF					
JOURNAL	EVAFPTAIGFELTGSNAGSNAMFIKLQMOQAQVNAVPMQYVTLQNTASSAVTIAC					

gene
SRT1GLYCNIPYRENEELLKTTLMIFGAVLLVVEVIVMYMLRN"
2132..2818
/locus_tag="BC1241"
2132..2818
/locus_tag="BC1241"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP08226.1"
/db_xref="GI:29894938"
/translation="MTLKNYGLKGTIYQSGIGCKTPHVOHLQGETGVDIRIAINY
KSQSYPEVLYEFASNNIKSEAIHILPLPREGTEIKNNKPRVALDYLRGNLFDNOMI
PLPEKAGADNDLNEKIERIKRAIEKAIYAFGERMGPENPDPSYFHKPENGCH
DHHMOQNEKMOQDDCIGMODGILLIHFEEKEEWIGLFAFOSQSMCTDEGHARVP
EHCYKGNK"
complement(2830..3663)
/locus_tag="BC1242"
complement(2830..3663)
/locus_tag="BC1242"
/codon_start=1
/transl_table=1
/product="integral membrane protein"
/protein_id="AAP08227.1"
/db_xref="GI:29894939"
/translation="MELVNSIFQILTSVIOQLSLIGVIVIGFLGLGLESITPMYMS
RAFGRKGFLLTAMIGVPIHELGHAIMCVLFRKIVATQFOFPTDSOGALGYVOYNO
KSVYORIGNEFGIGPIISGITALLIMRVPVPSYELFPTDSTASTINTEMIO
NMLSTFVLKSLFTIGNLNSPFLFLFIACISAHIALSKPIKSIDQVITMEFI
LFLFNITTGLPQYDSNQLIGVMKYNMILIAFSSVALLFSCISTLVSHSFYKIRGRNS
F"
3860..4630
/locus_tag="BC1243"
3860..4630
/locus_tag="BC1243"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP08228.1"
/db_xref="GI:29894940"
/translation="MKPSCHTGNAARAKPCGNGCHSTEEVAVSSGOEPEPEAKA
RPNETVQAKRFASGTFQFRKHALQATPAIMKSAIEIRKNCIVSLVLCIFGACIFR
MMSASAVTRTFARDISTPFEFGESITVEFFLLLTLEVGFIIVSGKMMKSSGFLE
VFGIMGTIAPTVIIVLVSFLSFLFLPILISVATYIGISITVAVLKLDGGLD
LVYLLIANVLIGATFMVMSYISTIIQALTKGTGF"
4650..5885
/locus_tag="BC1244"
4650..5885
/locus_tag="BC1244"
/codon_start=1
/transl_table=1
/product="IG hypothetical 16680"
/protein_id="AAP08229.1"
/db_xref="GI:29894941"
/translation="MNVCTKCGAQFEDGVQFCQNGSKRGPRVVKNNKSGTKVGTIL
LALFVIVIGLYLGASGYTQVAGVDRIATITLERDEKLAIEVTADPSVMTRESL
TLPFSYIKENPSYVNEELKJROGEKORDJIERDFSLTKDGKFFLEDFRKLAKTY
YTLLTNEKGYSLKMGKEIDKTDKKEKQYGLPGNOYFOSDKNDYKJLSREK
VVLKQSOYVNTIDITLOGOYITVOTNAPGATLYVNOKPVTAAGCERITGKPVATDS
ATTYLRNGESGRETTOYERTITLSSYNLPPQKSTEXTVYVNTPTPATGCHVNGFI
FPBSDIRKLTSADLTYSKEQLKARNEIYARHGHITQTKMQAVFAKQSWYRENPYF
KGTLDIETYNIELIKSRE"
complement(5931..6362)
/locus_tag="BC1245"
complement(5931..6362)
/locus_tag="BC1245"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP08230.1"
/db_xref="GI:29894942"
/translation="MCAGILALQLPDEQPEKRPAYSIDQESINATYISARAWLSH
DPTFQDIDIKATOPASASHACKSITFPVGSVKIINFSDDTITVTASPSDHALKIVPP
HSEALTSGLSHIQASTSGSSKVTFLFHIFSRDPGAP"

gene
6560..7894
/locus_tag="BC1246"
6560..7894
/locus_tag="BC1246"
/EC_number="1.6.99.3"
/EC_start=1
/transl_table=1
/product="NADH dehydrogenase"
/protein_id="AAP08231.1"
/db_xref="GI:29894943"
/translation="MNYVITIGDAAGSATNQIYTNDENANVTLEKEIYSVQOGL
PYVLSGALASTEKLIAHNVKTRDXYGIDAKVRIHEVYKVDTEKKMYVAEHRTKDVE
FPYRLIATVEKLVPMEMEGRELOGVILKTIPDAERIILKTETNKVEDYTIIIGGA
IGLMAEFVFLGKVMRIERNHDIGTLYDADMAEYIHKEDKHHIELTNNKAFK
GNERVEALETQDKGYKADLYVSVGVNPDPLGSTRNKRHKGAIEVNAVYQTNVOD
VYAGDCATPHVHVKELIHDHILPLCTTAKORSLAGLNLMDRRAFKGLTGCIITKPD
LTLARTGLNEKAEAGLHPYKTVKVDSTNMAGVYPNAKPLKLLIRSDTKQLGGV
IGEEGVKRIDVIAVALFNKMSIHDLBDVDYAPAPYNSVMDPIQQAARRAE"
8055..8465
/locus_tag="BC1247"
8055..8465
/locus_tag="BC1247"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP08232.1"
/db_xref="GI:29894944"
/translation="MALNMKTIEEMIASNARHEDRGHVVEEKVCGVGDNMTLIT
TKNVCFGRKVEVEFFQDHVIVIQEKDEYVEIKLKYDALITNSNLKTNKNTLLEK
FANGQINLDSLNDNYGTKMMLFARQIKSIFKL"
complement(8496..10037)
/locus_tag="BC1248"
complement(8496..10037)
/locus_tag="BC1248"
/EC_number="1.6.99.3"
/EC_start=1
/transl_table=1
/product="NADH oxidase (NOXASE)"
/protein_id="AAP08233.1"
/db_xref="GI:29894945"
/translation="WQLTFTLHHLHFSIDEITMPNHEVDKQAPLYKLYRNVPTPLS
LEIPLSNSSTPTPLNEIGHYLVWSEVYDQLCOLNSGRNVLRSISPGALVSNELY
ILKIDIDYPDGIIYHDAANHRLILLRGENFDSYADALGNCNHDGGAFTSMFV
KNFRKYNFESYRLOGDGVILIGOLLECAKQFYTGNGVYFOFLRALHHLGLSEGE
SVYAVIPSTSEPTNPMFNDVYHKNKTYSHLLQOIPLAHEHVRSHNDVYPMITK
INEASMIESTNHFOTLAHEHYQYHENAVALPKYBRSLSDYLQLCRKYSFADPTLT
KMDAELATLLQEASLSLPSNDLDGKITENAVASLIGCTYNAKDINGAVTYNSKT
HSIQPLRYGDLRYPLQSSMTDNNVNLFPVSLCLHVYGRKDYTNALGRYRIRHOMEA
GILVHKLVFAATYAMGMGHPPLSFDTNCSQDLYGIDAGNETSLIOPVGAYRARNMLK
GALHN"
complement(10053..11987)
/locus_tag="BC1249"
complement(10053..11987)
/locus_tag="BC1249"
/codon_start=1
/transl_table=1
/product="hypothetical Cytosolic Protein"
/protein_id="AAP08234.1"

Query Match 56.2%; Score 733.8; DB 1; Length 304680;
Best Local Similarity 72.6%; Pred. No. 5.3e-118;
Matches 948; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

DB 244018 ATGCCGAACCGAGTAATAGTAGCATAGTAGTCGCCGAACGATGGAAGAAATCTATATTTC 244077
1 ATGACTAACTATAGTAGCATATGTAGTAGGCTATATGTAGGAATCTATACATTTT 60
||||| 11111 11111 11111 11111 11111 11111 11111 11111
OY 61 AATAGATAGTTGGAGACGTGTTTCGATTGTTGAGATAGCGCAGTGTAAACAGAGAT 120
||||| 11111 11111 11111 11111 11111 11111 11111 11111
DB 244078 AATAGAAATTTGGAGAAAGATTTCATCGTAGAAGATATTCCAGGTGTAAACCGCGCAT 244137
OY 121 CGTATTATTCTTACAGTGAAGTGTAAACATGATTTCAATATTATTGATGACAGTGCT 180
||||| 11111 11111 11111 11111 11111 11111 11111 11111

Db	244138	GGTATTTATACGGCGTGAGAAATGGTTAAATCATGTAATTTAACTATTATGATACGGCGGA	2441197
QY	181	ATTGAAATTTGGTGAATGCACCATTTCCAAACACAAATTTAGACGCCAGGACAGAAATCCGCATA	240
Db	244198	ATTGGATTTTGGAGAGACGCCGGTTTGGACCAAAATTCGTCACAAAGCTACAGCTAGCAAAAT	244257
QY	241	GATGAAGCGGAGTGAATATTTTATTTGTAGCTAGCGCGGAAGGAAATTCACAAACCGCAT	300
Db	244258	GATGACGGGATGTTATATATTTTATTTATGACAAACCGCTCGTACCGGTGTAAACGACGCTGAT	244317
QY	301	GAATTGGTCGTCAAATTTTATATACAAATCTAAAAAACCGGTCGTATTTACGGTTAAACAA	360
Db	244318	GAGAGAGTTGCAAAAATTTTATTTATGCTTCAATTAACCGGTTGTACTTCAGTAATATA	244377
QY	361	GTAGATTAATATGGAATCGCTACAGACGCTATGATTTCTATTCATTTAGAGATTGGTGAA	420
Db	244378	GTAGCAATCCGGAAATCGGAGTACATTTTATGACTTTTATGACTTTTATGACTTTTGGCGAA	244437
QY	421	CCGTAATCCAAATCCAGGGTCACATGGTTAGGTCCTGTGGACACTTGTATATGCAGTTGT	480
Db	244438	CCATTTCCAAATTCAGGTACACACCGATTAAGGTTAGGTGACCTTCTAACAACAAGCTGCA	244497
QY	481	TCTCATTTGGTGAAGAGAGAAAGATCCCTTGTATGAAGATCAATTTGACATCATTT	540
Db	244498	CAACATTTTCCAAAGATTTGAAGAGACGGATATGATGAAGACACAAATTCGTTCTCTTTA	244557
QY	541	ATTGGAGACCAACGTAAGTAAATCAAGTTTATGTAATGCTATTTTATGTAAGTACATCC	600
Db	244558	ATTGGACGCTCAAAACGTAGGGAATCATCACTTGTAAATGCTCTCTTGGTCAAGCGCT	244617
QY	601	GTTATTCGTTTCTAATGTTGCAGGGAACAACGAGACGCTATGATACAGATATAGTTAT	660
Db	244618	GTAATTTGTAATATGTAAGTACAGGAACAACCTGTGATGCTGTGATACACCATATATATAA	244677
QY	661	GATGAGACAAGATTATGTTTAAATCATCTCGTGTATGCGTAAAAAAGAAAGATATAT	720
Db	244678	GATGAAAGATTTATGTAATTTATTTATTTATGATACAGCTGTATGCGTAAAAAAGAAAGCTAT	244737
QY	721	GAATCAACTGAAAAATTTCAATATTAAGCTTTAAAAACGATTGAACGTTCAATGTT	780
Db	244738	GAAAGTACAGAAAAATATATGTAATGTAATGCTGTGCGCTAAGACGATTGAACGTTCAAGCTT	244797
QY	781	GTTTATAGTGTTATTTGATGCGAAGAACAGGCATTATTGAACAAGATTAACGTTGTGCAGA	840
Db	244798	GTTTATAGCTGTTTATGATGCGAAGAAAGAAATTTATGACCAAGATTAATAAATAATCGCTGGA	244857
QY	841	TATGCACATGAACAAGTAAGACAGTCTGATTTGCTGTAATTAATTTGGATCTACTGGAA	900
Db	244858	TATGCTCTGATTTACAGGCGAGCTTGTATTGCTTTGTAATTAATTTGGATCAGTGAATA	244917
QY	901	AAAGATAGTAAAGCGATGAAGAAATTTGAAAGTGAAGTACGTAAGAAATTTCCAAATTTTA	960
Db	244918	AAAGACGAAAAACAATGAAGCAATTTGAAGAAATATCCGTCATTTCCAAATTCCTTA	244977
QY	961	GATTATGACAAATTTGCTTTGTGCTGTAAAGAACGACACAAGATTACGTACATTATTC	1020
Db	244978	GAATACGCAACAAATTTGTTTCTTATCTCAAAAACGAGAAAAAGTACACAAACATTAAATA	245037
QY	1021	CCTTACATTAATGAAGCAAGTGAACCAATATAAAAAGCTTCAAAAGTTCAACTTTAAAT	1080
Db	245038	CCGGTTATCGATGAGTAAATGAAGTCAATATATCTATCCGATTTCAAAACGAACGTAATTAAC	245097
QY	1081	GAACTCGTTACTGATGCAATTTCCATGTAACCTTACACCAACAGACAAAGAGTGAAGCTTG	1140
Db	245098	GATGAATTAATGAGTGGCGGTAGCGATGAATCGAGCGCAACATTAATTTGTAGCGCTCG	245157
QY	1141	AATGCTTTTATGCAACAAAGTTCATAGAACCCAGCGACATTTGTGTATTTGTTAAT	1200
Db	245158	AAATTTTCTATGCAACAAAGTTGGCTTAAACACCGACATTTGTGTGTTGTTAAAC	245217
QY	1201	GATGTAGATTAATGCAATTTTCTTATTAACGCTATTTAAGAAATCAATCCGTGCGCT	1260
Db	245218	GATCCGTAATTAATGCACTTCTCATATAGCGCTTCTTAAAGAAATTCGCTTATGTAATCA	245277

RESULT 9	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
1261	245278	TTTGGCTTTGAGGTACACCAATTCATTTACTCGCAAGAAGA 1305	AE017028	290117 bp	DNA	linear	BCT 30-APR-2003
AE017028		Bacillus anthracis str. Ames section 5 of 18 of the complete genome.	AE017028	AE016879			
LOCUS		AE017028.1 GI:30255149					
DEFINITION		Bacillus anthracis str. Ames					
ACCESSION		Bacillus anthracis str. Ames					
VERSION		Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus					
KEYWORDS		cereus group.					
SOURCE		1 (bases 1 to 290117)					
ORGANISM		Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Ostad, O., Helgason, E., Ristone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., Deboy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.					
REFERENCE		The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria					
JOURNAL		Nature 423 (6935), 81-86 (2003)					
MEDLINE		22608414					
PUBMED		12721629					
AUTHORS		2 (bases 1 to 290117)					
TITLE		Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Ostad, O., Helgason, E., Ristone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., Deboy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.					
TITLE		Direct Submission					
JOURNAL		Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA					
FEATURES		Location/Qualifiers					
SOURCE		1..290117					
		/organism="Bacillus anthracis str. Ames"					
		/mol_type="genomic DNA"					
		/strain="Ames"					
		/db_xref="taxon:198094"					
		331..1929					
		/locus_tag="Ba1197"					
		331..1929					
		/locus_tag="Ba1197"					
		/note="Similar to GP:10717143. and GP:10717143; identified by sequence similarity; putative"					
		/codon_start=1					
		/transl_table=1					
		/product="Oligopeptide ABC transporter, oligopeptide-binding protein"					
		/protein_id="AAP25159.1"					
		/db_xref="GI:30255150"					
		/translation="MMLGACSYKQDEPOANAKGDSGSAKQVINTLTVSEIPMTDTT LADDASSKVMNNTMEGLRYLKGDKLVYPVANSHTSSEDKKTFELRLDARKSNDD PVYAKDVMYAMORAVNPDTAKSGY IMLDVKNAEKIRLREISPOLQVAKAIDTFFEV ELNPVQYFLDIIVYPLFPLPLNKKYMEKDEKGLKENTLTYNPLNPLNKKHEOSQ LKNKPSWMDQKVEKLEPVNNVYKDGGAANLVNTSAIDVGSAAELVDYKKGGDKR TINDPQVFLFRNOKNADLANKNKIRAKISSAYIDBDGIGEVILNNGSGAGATGVPSD V KGDQDFFERENGLSKVDYKAEKKMEAKKELGDKRIEFLNPLNKKSGTGETVY					

gene
CDS
complement(1956..2099)
/locus_tag="BA1198"
/locus_tag="BA1198"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP25160.1"
/db_xref="GI:30255151"
/translation="MIGVITNSIGEKLEFGKEKKMKFVNSKSDVOSGMNVEGKGKGYKK
RTY"
gene
2731..3126
/locus_tag="BA1200"
2731..3126
/locus_tag="BA1200"
/note="similar to SP:P32687, PID:409796, GB:U00096,
PID:1790459, SP:P32687, PID:409796, GB:U00096, and
PID:1790459; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP25161.1"
/db_xref="GI:30255152"
/translation="MVTLYSPSCTSCSKAKLMEENHIPTERNIFSDPLTEIKE
ILNMTSGTDIISTRSKVFQELVNANESLPLODYKIMRDPGLRRPIMIDEKRLQ
VGVNEDIRRLPRTVRFQLREAROLVN"
complement(3176..3850)
/locus_tag="BA1201"
complement(3176..3850)
/locus_tag="BA1201"
/note="Identified by match to PFAM protein family HMM
PF03772"
/codon_start=1
/transl_table=11
/product="membrane protein, putative"
/protein_id="AAP25162.1"
/db_xref="GI:30255153"
/translation="MDLEFLTSLMIVGIDVILGDNALIVIALASRLPESKRNKAIL
IGTLAIIVRIITLAVYLDIPFLQIGVLLTLAVNLITNSNDLSSIQGTTL
FOAVRTIVADLVMGFDVNIAGAAHGRFSLVIGLISIPITLIGSKLILMIMERF
PELTYCGAIIATVAGKMTHEDERLATFEHNHNPSTAGIPYLFITLICIGFIYQOVR
LRNWKH"
gene
4112..4210
/locus_tag="BA1202"
4112..4210
/locus_tag="BA1202"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP25163.1"
/db_xref="GI:30255154"
/translation="MICHMSESLHNNKSGRGVVFWLKELMIIR"
4207..4860
/gene="mech"
/locus_tag="BA1203"
4207..4860
/gene="mech"
/locus_tag="BA1203"
/note="similar to GB:X03350, GB:X03350, GB:X03350,
GB:U00137, GB:M21692, GB:M24317, GB:M24308, GB:M24309,
GB:M24310, GB:M24311, GB:M24312, GB:M24313, GB:M24314,
GB:M24315, GB:M24316, GB:M37067, GB:X15447, GB:X15448,
GB:X15449, GB:X15450, GB:X15451, GB:X15452, GB:X15453,
GB:X15454, GB:X15455, SP:P00325, SP:P00326, SP:P07327,
PID:1223739, PID:178098, PID:178100, PID:178111,
PID:219428, PID:28416, PID:598143, PID:825622, GB:X03350,
GB:X03350, GB:X03350, GB:D00137, GB:M21692, GB:M24317,
GB:M24308, GB:M24309, GB:M24310, GB:M24311, GB:M24312,

gene
CDS
complement(1956..2099)
/locus_tag="BA1198"
/locus_tag="BA1198"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="negative regulator of competence Meca"
/protein_id="AAP25164.1"
/db_xref="GI:30255155"
/translation="MKRFITIDIDFDRGNREIYDRESPRELFWEMDEARDHDDF
FIDGFLIIOVAVDKGIEVLTKATLSKDGKLELPICVDKIIDPLDGEISLEQOE
LVEEVEOAGTNFEDGTFEFLIKFNDEVISLSHRIKFEDIKOELTSPEDRTIVYV
BFDEVLHDEEEDIRLISYLEVEGESTLTJHRSVEYKQIVKEHALETIRNNPAKT"
4933..6477
/gene="cls-2"
/locus_tag="BA1204"
4933..6477
/gene="cls-2"
/locus_tag="BA1204"
/note="similar to SP:066043, SP:066043, SP:066043,
SP:066043, SP:066043, and SP:066043; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase"
/protein_id="AAP25165.1"
/db_xref="GI:30255156"
/translation="MKNTLKIIFPVLILFALFVSRMFTDVAFYSGIDVSTLGI
ISLFTVSAFLGCIYFLENRRHPSKTLMLVLYGFPVPGFPAVILFEQNRKRMPO
KALLDEQAFLOYKGHEDYEEKILNNKHQELFLRLARGLANTISFOTERTLTNGD
ETPQAILDGLKRAKHHIMHEYIYVDKLTGETIKDILQSKSEGVVRRFLYDVGSEF
LSKSYIEELNDAGVEMIPFVRPFLINDKIYNNHRRIVIIDGNEGVFLDAGVSEF
LGKDYFGEFMRDTHLYLRGEAVOSLOLFLDOWFYMGEBVLAPEYLOAKAVEGSHMG
GVOUVAGGPDNKKWTIKHLYFAMTASAKSITMIAFPYLPDDLTLSALVALACTIDV
RLMPSKPKDKRTVYFASRSYFPELLDAVKRIEYKGLHSVYVYVDSDLASIGTAM
DMRSFLNFEVNAFLYDVSIRKLVDKRDLEESSEIHWDFHRRRLRRIVESTYR
LISPLL"
gene
6559..7803
/locus_tag="BA1205"
6559..7803
/locus_tag="BA1205"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP25166.1"
/db_xref="GI:30255157"
/translation="MIFPIAKRENGEKIHLIDHDELLHHMRRESFCTIACKEV
LMRLCKQRHMEHARKKVDSCLAFEAEISTYHMYGELLRYMRKQNVAVELEHYPEI
OORPDIETFERAGRKIAIEYOCANLSIEDLCKRTYVMOAGQIVIMIGNOLKHKHSA
WKKSESIAMFSLQSYPOPLIFEGCSKQSPKCTPTFSTSTISHTIYDIETTFE
EMLSRPVPOREVRVLRBKRRKDYRQNALPTLRNNHSLIBLYLOCCGNPNPFSOI
GVPLSPSFRQTNSTIQWFLIMKCTIGLAVGDCISLKYGCYIKRTIKRRMLPFSO
HIMKVAITEYMTFLCYAGLVKRVGWYTKRINDAVPLTGGVEIINHDEICLHNLSLF
EAKFMRCGCKGDIIDREGIT"
7854..9680
/gene="pepf-1"
Query Match 55.7%: Score 727.4, DB 1: Length 290117;
Best Local Similarity 72.3%: Pred. No. 6, 9e-117;
Matches 944: Conservative 0; Mismatches 361; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATGTAGCTTTGTAGGTAGGCGTAATGTAGTAATCTCAATTTT 60
Db 280374 ATGCCGAACCGATATGCAATAGTAGGCGCCCGAAGCGGAAACTCTACTTTTC 280433
QY 61 AATAGAAATAGTTGGAGACGTTGTCGATTTGTAAGATACCGCAGGTGTACAGAGAT 120
||||||| ||||||| | ||||| || ||||||| ||||||||| || |||||

Db 280434 AATAGATTGTTGGAGAGAGATTTCGAATCGTGAAGAAATATTCAGGTGAACGCCGAT 280493
 QY 121 CGTATTATCTTCAAGTGTGTTAACAACATGATTTCAATATTAATTAGAGGTGT 180
 Db 280494 CGTATTATAGGCGCTGAGAAATGTTAAATCATGATTAATTAACATTTATGATACAGGTGA 280553
 QY 181 ATTGAATGGTGATGACCATTCACAACAAATTAAGACGCGACAGAAATGCCATA 240
 Db 280554 ATTGATTTGGACGACGACATTTTGGACAAATTCGTCACAAACAGCTGAAGTAGCAATT 280613
 QY 241 GATGAGCGGATGTTTATTTATGTTAAGTGTGACGATTAAGCATTTGACACAAACGAT 300
 Db 280614 GATGAAGACAGATGTTATCATTTTATGACAAATGTCGTACGCTGTAAACAGCGGAT 280673
 QY 301 GAAATGTCGCTCAATTTTATCAATATCAAAACCGGTCGTTATGCGGTAAACAA 360
 Db 280674 GAAAGAGTTGCAAAATTTATTTATTCGTTCAAAATTAACCACTTCTACTGAGTAATAA 280733
 QY 361 GTAGATTAATGGAATTCGCTACAGACGTATGATTTTCTATTGATTAAGATTTGTGAA 420
 Db 280734 GTGGACAAATCCGGAATTCGCACTTATGATTTGATTCATTTAGGATTTGGGAG 280793
 QY 421 CCGTATCCAAATACAGGTGCACATGTTAGTCTGTGTCACCTTGTAGATGCACTTGT 480
 Db 280794 CCAATCCCAATCTCAGGTACACACGATTAAGATTAGTACTTATTAAGTGAAGCTCA 280853
 QY 481 TCATATTGTTGTAAGAGAGAAAGATCCCTTATGATGAAGATACAAATTCACATATCAT 540
 Db 280854 CAGCATTTTCCAAAGATTTGAAGAAAGCGATGATGATGAAGATACAAATTCCTTCTCTTA 280913
 QY 541 ATTGACGACCAAGATAGTAAATCAAGTTTATGTAATGCTATTTTATGTTAGTGAATCGC 600
 Db 280914 ATTGGCGCTCAAAATTAAGGAAATCATCTTGTAAATGCTCTCTTCTGCTCAAGACGT 280973
 QY 601 GTTATGTTTCTTAATGTTGAGGAGACAGAGAGAGCATTTATGATCAGATTTATGTTAT 660
 Db 280974 GTTATGTTTCTTAATGTTGAGGAGACAGAGAGCATTTATGATCAGATTTATGTTAT 660
 QY 661 GATGACCAAGATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 Db 281034 GATGCAAAAGATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 281093
 QY 721 GATCACTGAGAAATATTCAGTATTAAGACCTTTTAAAGCGATTGAACGTTCAAAATGTT 780
 Db 281094 GAAAGTACAGAAAGATTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281153
 QY 781 GTTATGTTGTTTATGATGACGAAAGAGCATTTATGAAACAAATTAAGCTTTGACGGA 840
 Db 281154 GTTATGTTGTTTATGATGACGTAAGAAAGCATTTATGAAACAAATTAAGCTTTGACGGA 281213
 QY 841 TATGACATGAACAGGTAAGAGCATGCTGATTTGCTGAATTAATGATGATGATGATGAT 900
 Db 281214 TACGCTCATGATTTCAAGAGAGCTGCTGATGCTGTTTAAACAAATGGGACGAGTGA 281273
 QY 901 AAAAGATGTAAGAAAGATTTGAAGATGAAGTACGTAAAGATTTCCAAATTTTAA 960
 Db 281274 AAAAGATGTAAGAAAGATTTGAAGATTTGAAGAAACATTTGCTGCTCAATTTTAA 281333
 QY 961 GATTATGACAAATTTCTTTGTTGCTGCTGAAGAAAGCAAGCATTTAGCTACATTTTC 1020
 Db 281334 GACTACGCAACGATTTTATTTATTTATGCAAAAACAGAAACGTCACAAACGTTAATA 281393
 QY 1021 CCTTACATTAATGAACAGTGAAGAAACATTAATAAAAGTGTTCAAATTTCAATTTAAAT 1080
 Db 281394 CCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 281453
 QY 1081 GAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 281454 GATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 281513
 QY 1141 AATGCTTTTATGACACAGTGTGATGAAGACCAAGCATTTTGTGATTTGTTAAT 1200
 Db 281514 AAGATTTTCTATGCAACAGATTTGCGTGAACCAACCAACATTTTGTGTTGTTGAAC 281573

QY 1201 GATGTAAGAAATTAATGATTTTCTTATTAAGACGATTTATGAGATCAATTCGTCGCT 1260
 Db 281574 GATCCAGAACTATTTGACCTCTCATATAGACGCGTCTTATAAATGCTTACGTGAATCA 281633
 QY 1261 TTTGCTTTGAGGTACACCAATTCATTTATGACTGGAAGAGA 1305
 Db 281634 TTTGCGCTTTGAGGAACGCCGATTCATTTATGCTGAGACAGA 281678
 RESULT 10
 AL591981/c
 LOCUS 347050 bp DNA linear BCF 06-JUN-2002
 DEFINITION *Listeria monocytogenes* strain EGD, complete genome, segment 9/12.
 ACCESSION AL591981
 VERSION AL591981.1 GI:16411141
 KEYWORDS
 SOURCE *Listeria monocytogenes*
 ORGANISM *Listeria monocytogenes*
 Bacteria; Firmicutes; Bacillales; *Listeriaceae*; *Listeria*.
 REFERENCE
 AUTHORS
 1
 Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A.,
 Baguer, F., Berche, P., Biocker, H., Brandt, P., Chakraborty, T.,
 Chardil, A., Cheouan, F., Couve, E., de Daruvar, A., Deloux, P.,
 Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L.,
 Dussurget, O., Enlihan, K.D., Esili, H., Portillo, F.G., Garrido, P.,
 Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Haut, J.,
 Jackson, D., Jones, L.M., Kaerst, U., Kretz, J., Kuhn, M., Kunst, F.,
 Kurapkat, G., Madueno, E., Maltounam, A., Vicente, J.M., Ng, E.,
 Nedjari, H., Nordstok, G., Novella, S., de Padlos, B., Perez-Diaz, J.C.,
 Pierrel, R., Rimmel, B., Rose, M., Schluteler, T., Sliemers, N.,
 Trierre, A., Vazquez-Boland, J.A., Voss, H., Wehlund, J. and Cossart, P.
 Comparative genomics of *Listeria* species
 Science 294 (5543), 849-852 (2001)
 JOURNAL 2153729
 MEDLINE
 PUBMED 11679669
 REFERENCE 2 (bases 1 to 347050)
 AUTHORS Glaser, P., Frangeul, L. and Rusniok, C.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
 Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
 Cedex 15, FRANCE
 E-mail: pglaser@pasteur.fr
 Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
 location/Qualifiers
 1. .347050
 /organism="Listeria monocytogenes"
 /mol_type="genomic DNA"
 /strain="EGD-e"
 /db_xref="taxon:1639"
 complement(95. .841)
 /gene="lmo168"
 complement(95. .841)
 /gene="lmo168"
 /note="lmo168"
 /note="similar to glucose 1-dehydrogenase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC9766.1"
 /db_xref="GI:16411142"
 /db_xref="SPTREMBL:O8YJ6"
 /translation="MNKVALVTGSSRGGLREIALALAKGYDIIVFSRRNRKAEKVQ
 OEIOJGRKRCVIFKANVGDVEXRELKAVDEFEGRDIFINNAAGVLRPLMELES
 HMDWTMNIKALLFAGOBARAKLMRHQSGKISLSSISIRYLENTYTVGVSKAIVE
 SLRYLAVELAPGIIVNVVSGLIETDNLNHPNREELLKDAVSKTPAGRMIEPDL
 VNAVFLASEKAMINGGITLVNGRTILY"
 complement(join(845. .1942,1948. .1953))
 /gene="lmo168"
 complement(845. .1942)
 /note="lmo168"
 /note="similar to A/G-specific adenine glycosylase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC9767.1"

/db_xref="GI:1641143"
 /db_xref="SPTREMBL:O8Y6J5"
 /translation="MRKMKRLTWDETKITAEQOALVSNYEANKRVLPHRENTPEYRIW
 VSEIMLOOTKVDVLPYFNREMTOPFPMESFENADOLIKMEGLGYSRANQOTA
 MKQVADPSGCEYDPLTITLSLKGVPPTAGAILSTAINQAEPAVDGNNVRIAYLE
 IGEDIMKASTKRIEEVEVLYQLDKRNPAAFNQGLMEIGALVCTPKPMCLPLDPC
 BAHKNGVETNPVKIKKYNKMKTELLSIVISEDKIAIEKRPENGLANMQOFTIE
 ISKKNDEVAKIQEFLHNGLEVLLEDEPTIAIKHVFSHLVMKMDIRVAKLQSAIPNEN
 WYFAEEEMKRLAFPVPYOKMMQAKDKFGE"
 /complement(850..855)
 /gene="Imo1688"
 /complement(850..855)
 /gene="Imo1688"
 /complement(1948..1953)
 /gene="Imo1689"
 /join(2129..3123,3130..3147)
 /gene="Imo1690"
 /complement(2129..2134)
 /gene="Imo1690"
 /complement(2143..3123)
 /gene="Imo1690"
 /note="similar to hypothetical proteins"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99768.1"
 /db_xref="GI:1641144"
 /db_xref="SPTREMBL:O8Y6J4"
 /translation="MDTGTHTVMGIALGLATVDVNVAGSSQAIGIMATITIGSOP
 DIDVLKLNNDYIRNHRGTHSLPLAIPMLISSITLYLPAPATFHLHLMFIA
 VGLHFDIFNAGYQAVRPEKEVWVAGFINTFMFIEGSHVVAIAAMLGSPVLP
 FVTLVIALYVAVRFTQRMKAHVNQLIPDSEIITASTHFQMRVAVTKDHY
 VGRARINISYEKEDRLPVPDNEIIRSAKKNLAAFISFSGVNMREEEDGTGY
 TETDLRYNSNGPVPVAVVKLDLKIYSSYGMIESETEKLYKKLAPYSI"
 /complement(join(3130..3147,3156..3617,3625..3630))
 /gene="Imo1691"
 /complement(3130..3147)
 /gene="Imo1690"
 /complement(3130..3147)
 /gene="Imo1691"
 /complement(3156..3617)
 /gene="Imo1691"
 /note="similar to deoxyuridine triphosphate
 nucleotidohydrolases"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99769.1"
 /db_xref="GI:1641145"
 /db_xref="SPTREMBL:O8Y6J3"
 /translation="MKVRGEVNVNASRKEPPEQTISLPTRDGKSGAGYDFSENQYTI
 VPGEKHFMTDIKSYMDEDEVINIVRSISGIKKGILLCNGTIDSSVSNPMDNGN
 IGIAIKNSNEPVSIEAGEBNAQGVFOKTLVADTDIVANESRGVGTGR"
 /complement(3625..3630)
 /gene="Imo1691"
 /complement(3629..3862)
 /gene="Imo1692"
 /complement(3629..3862)
 /gene="Imo1692"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99770.1"
 /db_xref="GI:1641146"
 /db_xref="SPTREMBL:O8Y6J2"
 /translation="MYTYLDELTAFLVKNPRLDKEQALWMTIEMLMSPDESSYAKAGY
 PYRGPEYAADYRQOIERHGSFLHOYERKDPK"
 /complement(join(3864..4673,4680..4685))
 /gene="Imo1693"
 /complement(3864..4673)
 /gene="Imo1693"
 /complement(3864..4673)
 /note="similar to hypothetical proteins"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99771.1"
 /db_xref="GI:1641147"

/db_xref="SPTREMBL:O8Y6J1"
 /translation="MKTISIVQOKNKERYNIFIDEKYNFSVDEEVLARYQLMKKVL
 TEADIEIKQADMVRKGLNKALNPLSHRBSREKIDYLRKQMEFPAIDELKKLAN
 MYINDFEAELEYTKOITTKLQPTIERBELVEGCTPEITISRYIEESDPAOLENA
 TKQAMTKMRKNKSAKMLQOKITIDLQKGTSELAKAAATEATSELDVABEALQ
 KOIEKTRNRKNKPKPIAKOKITISLMQGFSDYTIQSYLTENEISEEBEE"
 /complement(3872..3877)
 /gene="Imo1692"
 /complement(3872..3877)
 /gene="Imo1692"
 /complement(4680..4685)
 /gene="Imo1693"
 /join(4757..5673,5681..5708)
 /gene="Imo1694"
 /complement(4757..4762)
 /gene="Imo1694"
 /complement(4771..5673)
 /gene="Imo1694"
 /note="similar to CDP-abequose synthase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99772.1"
 /db_xref="GI:1641148"
 /db_xref="SPTREMBL:O8Y6J0"
 /translation="MHILLTGATGFIGDHLVHELEKSDHELYILTRQKLNRANVYI
 EMLNEDKLPLEDLPDVDCINLAGLIMEKWTYDRKRYIVNSRIEATSAISIVKM
 KSKPKIMNASAIQATYSSKSTYLDTEERTYADNDFLCTYVEMEKTAASALGIRV
 VYARRELVLTNGSGFRPVEKLFQYITGCRFRGROWYSHHYDDVVAIILFIEDHQ
 INGVNFIAPHVQEKFAERLCKKHKPKYKPPVKIILKIFLGERAMTILLOSQRPV
 EKLMSHFEEFRELQEDLDDL"
 /complement(join(5682..8291,8299..8304))
 /gene="Imo1694"
 /complement(5682..5709)
 /gene="Imo1695"
 /complement(5694..8291)
 /gene="Imo1695"
 /note="similar to putative membrane proteins"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99773.1"
 /db_xref="GI:1641149"
 /db_xref="SPTREMBL:O8Y6I9"
 /translation="MKERLMOAVAMFOKSTYVKYVTFVNAFVFEIINATGIDY
 PSIKENLTQSPEQITFMFTVGLIANTPMLLDIVYVKLLPKGFSFHHYASMTINT
 FNNIGGEGVLGASLNASFYGNKASHKELIALSKTALVLSGLSTYCLVSLATLLIP
 GFADHFNVMWMLAGLYPILFTTKKSKSLFYDLPIKRELTLIASLLEMFAF
 GCFALIGTLMGEPVDIFKVEPLFVIASVIGIASMVGGVGTFDVMILISLQVSOE
 LALAWMLFYRIFYIIPVAVGLLFEVQAKGKYNDELEGLPLFLQKVAHRLVIFYV
 GSGLLIILSSANPAIYHPELTKINPFLPQSOTTYAIFGLLGLGVARGIECKTKK
 AITIVIVLGCALFNTLAVFSKQALFGLIVLCLFLANEFYRKLVTYNSKVIID
 SIIFIVCLAGYIVIGYNSPNIKHSKEIDYLIASEHMLFMAEGEVLVSRYLIA
 IYLSITTEKELGSEFEAVKRYENHLAKNGEVSHTLRLKLFMAEGEVLVSRYLIA
 DKWVIMGEPTGNMCKMEALIEEVMADDEFGYVPPVYEVARGTWIPLYLHDGGEFRLG
 EGGFVQVONFTMSGKKKGERALMNKLEBEGTFEILIEPPNDNTWTTLRAVSDEVLQ
 GREKEGFSIGFRDYLYLBOAPLAIANKGCTVIGFASMPSTYDEMTSIDIMYSKEA
 PSGINDFLFTNLPEKAKEDGFQTFNAGMAPLANVGSKKAFUGERLAGLYRSQFV
 GFGKLNRFKSKYVTEWEQKFVAFRRKRSIAFTWLQMLILVKKRPLANSQVVLDFPLE
 EETKKPDSE"
 /complement(8299..8304)
 /gene="Imo1695"
 /complement(8318..9007)
 /gene="Imo1696"
 /complement(8318..8992)
 /gene="Imo1696"

Query Match 52.4% Score 683.2; DB 1; Length 347050;
 Best Local Similarity 70.2%; Pred. No. 2.9e-109;
 Matches 916; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 QY 1 ATGACTAACTATAGTACTATTTAGTAGAGCCTAATGTAGTAAGTAACTTACAAATTTT 60
 III I IIIII I IIIII IIIII I I I I I I I I I I IIIII I IIIII I

Dp	259438	ATGGCAAAACAGTGTGTAGGCATGTGTGGACGTCCAAACGTTGGCAATTCGCAATATTTT	259379
Qy	61	AATGAATAGTTGGAGAACGTGTTTCGATTGTTGAAGATAGCCAGGTGTAAACAGAGAT	120
Dp	259378	AACGAAATCGTTGTGGAACGTGTTCATAGTGAAGATGTTCGCGTGTGACACGTGAC	259319
Qy	121	CGTATTATTCCTCAAGGTGAGTGTGAACACATGATTTCAATATATATTTGATACAGTGGT	180
Dp	259318	CGCATATATTAATTCAGCGGAATGCGTTGGAAAAAGAAATTTAAACATTAATTAATCAGGTGGT	259259
Qy	181	ATTGAAATTTGGTGAGCACCATTCACAAACACAAATTAGAGCCAGCACAAATTCGCGCAT	240
Dp	259258	ATTGATCTTTCCGATGAACCATTTTAAAGCAAAATTCGCGCACAAACGGAAATTCGAATT	259199
Qy	241	GATGAGCGGATGTTATTTATTTATGTTAACTGCGGTGAAGATTGACACAAACGAT	300
Dp	259198	GATGAAACGACAGCATTAATATTTTATATACCAATGTGCGAAGAGATACCGATGACAC	259139
Qy	301	GAAATGTCGCTCAAAATTTTATACAAATCTAAAAAACCGGTGTGATTACCGTTAAACAA	360
Dp	259138	GAACAAAGTAGCAAAAAATTCCTTACGCGTCTAATPAAACCAATGTGTTTACCGATTATPAAA	259079
Qy	361	GTAGATTAATATGAAGATCGTACACAGCGTGTGATTTCTATTCATTTAGATTGGTGA	420
Dp	259078	GTAGATTAACCCAGAAATCGAGATCAAAATTTTGACTTTTATTTCTTGTGGTTGGTGA	259019
Qy	421	CCGTAATCCAAATATCAGGCTCACATGTTTAAAGTCTGTGTGACTTGTTAGATCGAGTTGT	480
Dp	259018	CCGTAATCCAAATTTCTGTTCCATGACATGAGAGGCTTGGCGACATGCTGATCTGTTCG	258959
Qy	481	TCTCATTTTGGTGAAGAGAGAAACATCCTTATGATGAAGATPACAATTCGACTATTCAT	540
Dp	258958	GCTCATTTTCCAAAAAGAAAGAGAGAAAGAAATPCCAGATGACACAGTGAATTCAGTTTA	258899
Qy	541	ATTGAGACACCAAACGTAGTAAATCAAGTTTACTAATATGCTATTTTACGTAAAGATCG	600
Dp	258898	ATCGGTGACCAAAATGTTGATTAATCTTCTATTTCTAATAGCACTGTTGGAGAAATGCT	258839
Qy	601	GTTATCGTTTCTAATGTTTGACAGGACAAACGAGACGCTATTGATPACAGATATAGTAT	660
Dp	258838	GTTATGTTTTCAGATTAATTTGCGGGTCAACCTCGTATGCAATGATTAACATTTACTTTC	258779
Qy	661	GATGACACAAGTATGTTTAAATCATCTGCTGTATGCGTATGCGTAAAAAGAAAAGTATAT	720
Dp	258778	GATGGCCAGAGTTATGATCATGATGATATACAGCCGGAATGAGAAAAACGTGGAAAGTAT	258719
Qy	721	GAATCAACATGAGAAATATTTACGTATTAGACGCTTTAAAAACATTGAACGTTCAATGTT	780
Dp	258718	GAAGACACAGAAATATATGTTTACGTGCAATGACAGCAATTTGAACGCTCGCACGTT	258659
Qy	781	GTTTATGTTGTTATGATGACAGAACAGCATTTATTGAACAAGATAAACGTTGTGACGA	840
Dp	258658	GTTCTTGGTTATTCACAGCGAAGAAAGATATTCGTGACCAAGATTAACGAATGTGTGA	258599
Qy	841	TATGCACATGAAACAAGTAAAGCACTCGTATTTGTCGTAAATTAATTTGGATACTGTGAA	900
Dp	258598	TATGGCGATGATGCCGGACGTGCAATTTATTAATGTTGGAAACAAATTTGGATCAATTTAC	258539
Qy	901	AAAGATAGTAAACGATGAGAAAAATTTAAAGATGAGATGAGTAAAGAAATTCACATTTTAA	960
Dp	258538	AAAGATGAAAAACAATTAATGATATGACACAGAAAGTATTCGTGACCAATTTCCAATTTCTTA	258479
Qy	961	GATTATGACAAATTTGCTTTTGTGTCTGTAAAGACGACAAAGATTAAGTCATTTATTC	1020
Dp	258478	AGTATATGACCAAAATGTTTGTGTATCTGTAAACAAAACAGCGTTAAATATATCATTC	258419
Qy	1021	CCTTACATTAATGAAGCAAGTGAAGAAACATAAAAAACGTGTTCAAAAGTTCAACTTTAAT	1080
Dp	258418	CCACTTATTAACCAAGTAAAGGATTAACCATTCATTAATTCGTGTACATCAAGTATGCGAAT	258359
Qy	1081	GAACTGCTTACTGATGCAATTTTCCATGAACCTTAACCAACAGACAAAGGTAGACGTTTG	1140
Dp	258358	GATGTATTAATGATGATGCTGTTCATGATGAACCATATCCATATGATGAATGAAGTTAAAGTAC	258299

QY	11411	AAATGCTTTTATGCAACAACTGGCTATAGAACCCACGACATTTGCTGATTTGTAAT	1200
Db	258298	AAAGATTTTCTATACACACCAATGGCGGTAAACCCACCAAGCTTTTGATTTGTAAT	2582395
QY	1201	GATGTAGAATTAATGCAATTTTCTTATTAACGCTATTATAGAGAAATCCGTCGCCGT	1260
Db	258238	GATCCAGAACTCATGCAATTTTTCGATTAAGAACGTTTCTTAGAAACCGGATTAAGAAAGCA	258179
QY	1261	TTTGCTTTTGAAGGTACACCAATTCATATTATAGCTCGAAAGAG	1304
Db	258178	TTTCCATTTTGAGGTAACGCCAATTCGAGTAATATGCTCGTAAGCG	258135
RESULT 11			
AX641670/c	AX641670	349980 bp	DNA
LOCUS	Sequence	2860 from Patent WO0101118.	linear
DEFINITION	AX641670		PAT 21-FEB-2003
ACCESSION	AX641670		
VERSION	AX641670.1	GI:28474431	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE:			
AUTHORS			
1	Buchrieser, C., Frangeul, L., Couve, E., Rusanok, C., Esht, H., Dehoux, P., Dussurget, O., Chetoui, F., Nedjari, H., Glaser, P., Kunst, F., Cossart, P., Daniels, J., Goebel, W., Kreft, J., Kuhn, M., Ng, E., Vaquer, B., Amend, A., Chakraborty, T., Doman, E., Hahn, T., Tierrez-Martinez, A., Durand, L., Perez-Diaz, J. C., Baquero, F., Berche, P., Charbit, A., Durand, L., Perez-Diaz, J. C., Baquero, F., Garcia del Portillo, F., Gomez-Lopez, N., Madueno, E., de Pablos, B., Wehland, J., Kaerst, U., Entian, K. D., Hauf, J., Rose, M., and Voss, H.		
TITLE			
JOURNAL			
FEATURES			
source			
1.	349980		
/organism="Listeria monocytogenes"			
/mol_type="genomic DNA"			
/db_xref="taxon:1639"			
/note="Seq 1 for big: original length: 2944528-splitted into 10 <23>seq 0001 : From 0.000.001 to 0.349.980 length: 349.980 <23>seq 2855 : From 0.300.001 to 0.649.980 length: 349.980 <23>seq 2856 : From 0.600.001 to 0.949.980 length: 349.980 <23>seq 2857 : From 0.900.001 to 1.249.980 length: 349.980 <23>seq 2858 : From 1.200.001 to 1.549.980 length: 349.980 <23>seq 2859 : From 1.500.001 to 1.849.980 length: 349.980 <23>seq 2860 : From 1.800.001 to 2.149.980 length: 349.980 <23>seq 2861 : From 2.100.001 to 2.449.980 length: 349.980 <23>seq 2862 : From 2.400.001 to 2.749.980 length: 349.980 <23>seq 2863 : From 2.700.001 to 2.944.528 length: 244.528"			
BASE COUNT	99761 a	75437 c	61177 g
ORIGIN			
Query Match	52.4%	Score 683.2:	DB 6:
Best Local Similarity	70.2%	Pred. No. 2.9e-109:	Length 349980:
Matches	916:	Conservative	0:
		Mismatches	388:
		Indels	0:
		Gaps	0:
QY	1	ATGACTAAACCTATAGTAGCTATTTAGGTAGGAGCTTAATGTAGTAAATCTACAAATTTT	60
Db	212438	ATGCGAAACACGAGTGTGAGCAATTTTCGAGAGCTCCAAAGCTTGGCAATCGACTATTTT	212379
QY	61	AATAGCAATAGTTGGAACAGTGTTCGATTTGTAAGATACGCCAGGTGTAAACAGAGAT	120
Db	212378	AACAGAAACGTTGTGGAACAGTGTTCGATTTGTAAGATACGCCAGGTGTAAACAGAGAT	212319
QY	121	CGTATTTATTTCTTCAGGTGAGTGTAAACACATGATTTCAATATTATTTATACAGTGT	180
Db	212318	CGCATTTATTTATTTACGCGGAATGCTTGGAAAAAGATTTTAAACATTTATGATACAGTGT	212259

Oy	181	ATGAAATGGTGATGACCAATTC	240
Db	21258	ATTGATCTTCCGATGAACATTTT	212199
Oy	241	GATGAAGCGGATGTTATTTTAT	300
Db	212198	GATGAAGGAGACGTAATTTAT	212139
Oy	301	GAATGCGCTCAATTTTATAC	360
Db	212138	GAACAGTAGCAAAAAATTC	212079
Oy	361	GTAGATAATTAAGAAATCG	420
Db	212078	GTAGATAATTAAGAAATCG	212019
Oy	421	CCGATCCAAATTCAGGCTC	480
Db	212018	CCGATCCAAATTTCTGGTTC	211959
Oy	481	TCTCATTTTGGTGAAGAG	540
Db	211958	GCTCATTTTCCAAAAAGAG	211899
Oy	541	ATTGACGACCAAACTAGT	600
Db	211898	ATCGGTGACCAAAATGTTG	211839
Oy	601	GTTATCGTTCTAATGTTG	660
Db	211838	GTTATGTTTGAATATGTC	211779
Oy	661	GATGACCAAGATTAATG	720
Db	211778	GATGACCAAGATTAATG	211719
Oy	721	GAATCACTGAGAAATAT	780
Db	211718	GAACACAGAGAAATAT	211659
Oy	781	GTTTATAGTGTTATGAT	840
Db	211658	GTTTGTGTTATCAACG	211599
Oy	841	TATGACATGACAAAGT	900
Db	211598	TATGGCATGATGCCGAC	211539
Oy	901	AAAGATAGTAAACGAT	960
Db	211538	AAAGATGAAAAAATTA	211479
Oy	961	GATATGACAAATTCG	1020
Db	211478	AGCATATGACCAATG	211419
Oy	1021	CCTTACATTAATGAAG	1080
Db	211418	CCACTATTAATCAAG	211359
Oy	1081	GAAGTGGTACTGATG	1140
Db	211358	GATGTATTAAGTAG	211299
Oy	1141	AATGCTTTTATGAC	1200
Db	211298	AAGATTTTCTATAC	211239
Oy	1201	GATGATGAATTAATG	1260
Db	211238	GATCGAAGATCATG	211179
Oy	1261	TTTGGTTTGAAGTAC	1304

Db	211178	TTTCCATTTGAGGTACGCAAT	211135
RESULT 12	AL596170/c		
LOCUS	AL596170	313450 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Listeria innocua C11p11262 complete genome, segment 8/12.		
ACCESSION	AL596170		
VERSION	AL596170.1	GI:16414292	
KEYWORDS			
SOURCE	Listeria innocua		
ORGANISM	Listeria innocua		
REFERENCE	1		
AUTHORS	Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloeker, H., Brandt, P., Charraorty, T., Charbil, A., Chetoui, F., Couve, E., de Darvar, A., Denoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fajl, H., Portillo, F.G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kref, J., Kuhn, M., Kunst, F., Kurapk, G., Madueno, E., Maitournam, A., Vicente, J.M., Ng, E., Nedjari, H., Nordiek, G., Novella, S., de Pablo, B., Perez-Diaz, J.C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehlend, J. and Cossart, P. Comparative genomics of Listeria species Science 294 (5543), 849-852 (2001)		
JOURNAL	21537279		
MEDLINE	11679669		
PUBMED	2 (bases 1 to 313450)		
REFERENCE	Glaser, P., Frangeul, L. and Rusniok, C.		
AUTHORS	Submitted (09-Jul-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE		
COMMENT	E-mail: pglaser@pasteur.fr		
FEATURES	Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.		
SOURCE	1.313450		
location/Qualifiers			
gene	/organism="Listeria innocua"		
CDS	/mol_type="genomic DNA"		
	/strain="C11p11262"		
	/db_xref="taxon:1642"		
	complement(287..1444)		
	/gene="lin1790"		
	complement(287..1444)		
	/note="similar to transmembrane transport proteins"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAC97021.1"		
	/db_xref="GI:16414293"		
	/db_xref="SPTREMBL:Q92AX8"		
	/translation="MGAKGFWILTMVAISGLSGVLPLPIAILLEKGVASAGINGF		
	HATGIYGLVLSPEIEALPHKYGKPIILVGGGLVAIAIFAPIMFNLVFWLRL		
	IGIGDHLMHFSQWIGAMSDPSKGRNIAIYGLFVSGFAIPQDVLAKANVLP		
	FLSGIIVLIAMSLVWFINDPGEKAVIRKISFKSIRFSGVFMAYAMIPPLYG		
	ILERTGLNATFPVYGLRDGLDPMIAIMIISSPSVGTITFOVPYIGIISDKFRKMLPL		
	TEGAAVVEVLNAEVRIPLEVFIEFFITGLISLVSIGLSTYMDLPLPLPAGNIVL		
	GMFSLGIIGPSATGMIGIFGNQIFVFAVAGILLCGLLALGGRKRAH"		
	complement(1508..1579)		
	/product="rRNA-Glu"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 69.21"		
	complement(1585..1675)		
	/product="rRNA-Ser"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 63.76"		
	complement(1712..1784)		
	/product="rRNA-Asn"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 85.46"		
	complement(1791..1864)		
	/product="rRNA-Ile"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 92.91"		

QY	121	CGATTATATCTCAAGGAGGCTGTTACACATGATTTCAATATTAATTTACACAGCTGT	180
Db	276690	CGCAATATTAATTCACGGAAAGCGCTTGAAAAAATTTAACATATTAATGATACAGCTGT	276631
QY	181	ATTGAAATGGGATGATGCACATTTCCAAACCAAAATTAGAGCGCAGGACAGAAATCGCATTA	240
Db	276630	ATTGATCTTTCCGACGAACCATCTTTAGACCAAAATTCGGCACACAGCGAAATCGCAAT	276571
QY	241	GATGAAGCGGATGTATTAATTTTATATGTTTAAACGTCGCGGAAGATTGACACAAACGAT	300
Db	276570	GATGAAGCAGACGTAATTAATTTATTAACCATGTGCGTGAAGGGGTTACCGATGCAGAC	276511
QY	301	GAATGTCGCTCCAATTTTATACAAATCTAAAAAACCGGCTGATTTAGCGGTTAACAA	360
Db	276510	GAACAAGTAGCAAAAATTTCTTAACCGGCTAATTAACCAATGTGTTTACGATTAATAAA	276451
QY	361	GTAGATTAATATGGAATAGCGTACAGCGGTATGATTTCTATATTCATTAGATTTGGTGAA	420
Db	276450	GTAGATTAACCCAGAAATGCGTATCAGATTTATGACTTTTATCTTCTGGTTGGTGAG	276391
QY	421	CCGATATCCAAATATCAGGGCTACATGGTTAGGCTCTTGAGCTTGTATGATGCAATGT	480
Db	276390	CCGATATCCAAATTTGCTTCATCATGAGACTAGGGCTTGATGATGCTGATGCTGCT	276331
QY	481	TCTCATTTTGGTGAAGAGAGAGAGATCCTTATGATGAAGATACAAATTCGACTATTCAT	540
Db	276330	GCTCATTTTCCAAAAGAAAGAGAGAGAGATCCAGATGAAGAACATTAATTTAGTTTG	276271
QY	541	ATTGAGACACCAAGCTAGTAAATCAACAGTTTAGTAATAGCTATTTAGGTGAAGATCCG	600
Db	276270	ATTGATCGGCCAAATGTTGGTAAATCATCTATTTAAACGCACTTCTTGCGCAAGACCG	276211
QY	601	GTTATCGTTTAAATGTTGACAGGACCAACGAGACGCTATTGATACAGATTAAGTAT	660
Db	276210	GTCATGTTTTCGACATTTGCGGGAACCTACTCGTGATGCAATTTGATCAACTTATACGTT	276151
QY	661	GATGACACAAGATTATGTTTAAATCGATACTGCTGTATGCGTAAAAAGAAAAGTATAT	720
Db	276150	GATGCTCAGAGATTATGTCATGATTAATACAGCCGGAATGAGAAACGTGGAAAGGTAT	276091
QY	721	GAATCACTGGAATATTTCACTAATTAAGCTTTAAAGCGATTGAACGTTCAATGTT	780
Db	276090	GAAGACACAGAAATATATGTTTACTGCAATGAGAGCAATTAAGCTCCGATGTT	276031
QY	781	GTTTATAGTATATGATGACAGAACGCAATATTAAGACAGATTAACGCTGTGAGG	840
Db	276030	GTTCTGTGGTTATCAACGCGAAGAAAGATTTCTGAGACAGATTAAGCGGATTCGTGGA	275971
QY	841	TATGCACATGAACAAGGTAAAGACAGTCTGATTGTCTGAATTAATGAATGGAATACTGTGAA	900
Db	275970	TATGCGCATGATGATCCGAGCGGCTATCATTTATTTAGTAGAACAAATGGGATGCAATTAAC	275911
QY	901	AAAGTATATAAACCATGAAGAATTTGAAGATGAAGTACGTAAAGATTTCCATTTTA	960
Db	275910	AAAGATGAAGAAACAATTAATGTATGAGACAGAAATATTCGGGAGACAGTTCCAAATCTTA	275851
QY	961	GATATGACAAATGCTTTTGTCTGCTATAAGAACGCAACAAGATTAGCTACATTAATTC	1020
Db	275850	AGCTATGCACCAATGTTTCTGTATCTGCTAAACAAACAAACGCTTAAACAACATTAATTC	275791
QY	1021	CCTTACATTAATGAAGCAAGTGAAACCATTAACAAACGCTTCAAAATTCACATTTAAAT	1080
Db	275790	CCGCTCATTAACCAAGTAAGCATATCATTTACATTAACGGGTACAAATCTAGTATGTAAT	275731
QY	1081	GAAGTCGTTACGTAGCAATTTCCATGAACCTTACACCAAGACAGACAAGGTAGACGTTG	1140
Db	275730	GATGTTATTAAGATGCGGTTGCCAATGATTCATCCACCAATGGATTAAGGTAAACACAT	275671
QY	1141	AATGCTTTTATGAACAACAGTGTCTATAGAACCCAGCAATTTTGTGATTTGTAAT	1200
Db	275670	AAATATTTCTATACACACACAGTGGCTGTTAAACGCCCAACATTTTGTGATTTGTAAT	275611
QY	1201	GATGTAGATTAATGCAATTTTCTTATAAACGCTATTTAGAGATCACTAAATCCGTGCCGT	1260

Db	275610	GATCCAGAACTAATGATTTCTTATGAAACGGTTCTTATGAAAAACCGGATTAAGAGA	275551
Qy	1261	TTTGGTTTGAAGTACACCAATTCATATATAGTCGAAAG	1304
Db	275550	TTTCCGTTGATGTACGCCAATTCAGTAATTTGCTGTAAGCG	275507
RESULT 13			
AX413016	319630 bp	DNA	linear
LOCUS	AX413016	Sequence 7 from Patent WO0228891.	PAT 02-SEP-2002
DEFINITION	AX413016	GI:21445474	
ACCESSION	AX413016.1		
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT	105207 a 55428 c 66726 g 92263 t	6 others	
ORIGIN			
Query Match	51.9%;	Score 676.8;	DB 6; Length 319630;
Best Local Similarity	69.9%;	Pred. No. 3.8e-108;	
Matches	912;	Conservative 0;	Mismatches 392; Indels 0; Gaps 0;
Qy	1	ATGACTTAACCTATAGTACTATTTAGGTAGGCGCTAATGTAGTAATCTACAATTTT	60
Db	50618	ATGGCAAAACCAAGTTAGTCAGTATGTCGGACGTCGCAACGTTGGCAAAATCGACTATTTT	50677
Qy	61	AATGAATAGTGTGGAGAACGCTGTTGCGATTTGGAAGATACGCCAGTGTACACGAGAT	120
Db	50678	AACAGATAGTGTGTGTAACGCTGTTCCATAGTGAAGATGTTCCGGGTGACACGAGAC	50737
Qy	121	CGATTTATTTCTTCAGGTAGTGGTGAACACATGATTTCAATATTTATGTACAGNGGT	180
Db	50738	CCCATATTAATTCACGCGAATGCTTGGAAAAAATTTTACATTTATGTACAGNGGT	50797
Qy	181	ATTGAATTTGGTATGCACCATTTCCAAACACAATTAAGACCGCAGCAGAAATCGCCATA	240
Db	50798	ATTGATCTTCCGACAAACCATTTCTTAAGCAAAATTCGGGCACAAAGGGAATTCGCAATT	50857
Qy	241	GATGAAGCGGATGTTATTTATTTTATGTTTAACGTCGCTGAAGGATTTGACACAAGCAT	300
Db	50858	GATGAAGCGAGCTAATTTATTTTATTAACCAATGTCGTGAAGGGTTTACCGATGCAGAC	50917
Qy	301	GAAGAGGTGCTCAATTTTATCAATATTAATAAAACCGGTCGTATTAGCGTTAAACA	360
Db	50918	GAAACAAGTACAAAAATTTCTTACCGGCTCAATTAATAACCAATTTTATAGCGATTAATA	50977
Qy	361	GTAGATTAATATGGAATGCGTACAGACGCTATAGATTTTCAATTCATTTAGGATTTGTA	420
Db	50978	GTAATATTAACCAAAATGCGTATGATTTATGACATTTTATTTCTGTGGTTGGTGAG	51037
Qy	421	CCGATATCAATATCAGGCTACATGCTTTAGTCTTTGGTGATCTTTGATGCAAGTTGT	480
Db	51038	CCGATATCAATTTCTGCTCTCATGACCTAAGGCGTTGCTGATATGCTGTGCTGTTCT	51097
Qy	481	TCTCATTTTGGGAAGAGAAGAAATCCATTATGATGAAGATCAATTCAGATTCATAT	540
Db	51098	GCTCATTTTCCAAAGAGAAGAGGAAGATACCCAGATGAACAGTGAATTTAGTTTG	51157
Qy	541	ATTGGACGACCAACCTAGTAATCAAGTTTATGTAATTCATTTTATAGTGAAGATCGC	600

Db 51158 ATTGGTCGGCAATGTGGTAAATCATCTATTCTTAACCCACTTCTTGGCGAAGACCG 51217
QY 601 GTTATGCTTTCTAATGTTTGCAGGACAAACAGAGACGCTATTGATACAGATATAGTTAT 660
Db 51218 GTCATTTGTTCTGACATTTGGGGAACTACTCGTATGATCAATTTGATCACTTATAGCTTT 51277
QY 661 GATGACAGATATATGTTTAAATCAGTACTGCTGATGCTGATGCTGATGCTGATGCTGAT 720
Db 51278 GATGTCAGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51337
QY 721 GAATCACTGAGAAATATTCAGTATTAGAGCTTTTAAACCGATTTGAGCTTCAATATGTT 780
Db 51338 GAAGACAGAGAAATATATGATGTTTTCATGATGATGATGATGATGATGATGATGATGAT 51397
QY 781 GTTTAGTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 51398 GTTCTGTGTTTATCAACGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 51457
QY 841 TATGACATGACAGATTAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 51458 TATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51517
QY 901 AAGATAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 51518 AAGATAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51577
QY 961 GATTATGACAAATGCTTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 51578 AGCTATGACAAATGCTTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 51637
QY 1021 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 51638 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51697
QY 1081 GAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 51698 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51757
QY 1141 AATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 51758 AATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51817
QY 1201 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 51818 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51877
QY 1261 TTTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1304
Db 51878 TTTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51921

RESULT 14
AX417046/c 349980 bp DNA linear PAT 02-SEP-2002
LOCUS Sequence 4037 from Patent W00228891.
DEFINITION AX417046
ACCESSION AX417046 GI:21449656
VERSION
KEYWORDS
SOURCE Listeria innocua
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

REFERENCE
1 Kunst, F. and Glaser, P.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 4037 11-APR-2002;
INSTITUT PASTEREUR (FR) : CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
LOCATION/Qualifiers
1. 349980
/organism="Listeria innocua"
/mol_type="genomic DNA"
/db_xref="taxon:1642"

/note="seq 2058, original length: 3,011,208 replaced
by-seq 2058: 0.000,001 to 0.349,980-seq 4032: 0.300,001 to
0.649,980-seq 4033: 0.600,001 to 0.949,980-seq 4034:
0.900,001 to 1.249,980-seq 4035: 1.200,001 to
1.549,980-seq 4036: 1.500,001 to 1.849,980-seq 4037:
1.800,001 to 2.149,980-seq 4038: 2.100,001 to
2.449,980-seq 4039: 2.400,001 to 2.749,980-seq 4040:
2.700,001 to 3.049,980-seq 4041: 3.000,001 to 3.011,208"
BASE COUNT 101055 a 72969 c 60688 g 115268 t
ORIGIN

Query Match 51.9% Score 676.8; DB 6; Length 349980;
Best Local Similarity 69.9% Pred. No. 3,7e-108;
Matches 912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;

QY 1 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 26680 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266751
QY 61 AATGAT 120
Db 266750 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266691
QY 121 CGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 266690 CGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266631
QY 181 ATGAT 240
Db 266630 ATGAT 266571
QY 241 GAT 300
Db 266570 GAT 266511
QY 301 GAATGAT 360
Db 266510 GAATGAT 266451
QY 361 GTATGAT 420
Db 266450 GTATGAT 266391
QY 421 CCTATGAT 480
Db 266390 CCTATGAT 266331
QY 481 TCTATGAT 540
Db 266330 TCTATGAT 266271
QY 541 ATTGAT 600
Db 266270 ATTGAT 266211
QY 601 GTTATGAT 660
Db 266210 GTTATGAT 266151
QY 661 GAT 720
Db 266150 GAT 266091
QY 721 GAATCACTGAGAAATATTCAGTATTAGAGCTTTTAAACCGATTTGAGCTTCAATATGTT 780
Db 266090 GAATCACTGAGAAATATTCAGTATTAGAGCTTTTAAACCGATTTGAGCTTCAATATGTT 266031
QY 781 GTTTAGTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 266030 GTTTAGTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 265971
QY 841 TATGACATGAT 900
Db 265970 TATGACATGAT


```

Db      265970 TATGCGCATGATGCCGAGCCGCTATCATTTATGTAGTGAACAATGGATGCAATTAAC 265911
QY      901 AAGATAGTAAACGATGAAGAAATTTGAGATGACGATGAAGATTCACATTTTGA 960
Db      265910 AAGATGAAAAAACATTAATGATGTGACAGAGATATTTGCGGACGATTCACATTTCTTA 265851
QY      961 GATTATGACAAATTTGCTTTGTGCTGCTAAAGACGACACAGATTACGTACATTAATTC 1020
Db      265850 AGCATGACACCAATTTGTTTGTGCTATGCTGCTAAACAAACACAGCTTAACACATTTATTC 265791
QY      1021 CCTTACTTTAATGACGACAGTGAACCATTAACCAACGCTGTTCAAGTTCACTTAAT 1080
Db      265790 CCGCTCAATTAACCAAGTAAACGATTAATCATTTACGCGTACATCTAGTATCTAAT 265731
QY      1081 GAAGTCGTTACTGATGCAATTTCCATGACCCCTACACACAGCAAAAGTAAAGCTTGG 1140
Db      265730 GATGTTATTTAGTATGCGGTTTGCATTAATCCATACCAATGATTAAGTAAACGACTT 265671
QY      1141 AATGCTTTTATGCAACACAAAGTTGCTATAGAACACGACATTTGTTATTTGTTAAT 1200
Db      265670 AAATATTTCTATACACACAAAGTGGCTGTTAAACGCGACATTTGTTGTTATTTGTTAAT 265611
QY      1201 GATGTAGAAATTAATGCAATTTTCTTATTAACGCTATTTAGAAATCAATCCGCGCT 1260
Db      265610 GATCCAGAACTAATGCAATTTCTTATGAACGTTTCTTAGAAACCGGATTAAGAGAACA 265551
QY      1261 TTTCGTTTGAAGTACCAATTCATATTATAGCTGGAAGAG 1304
Db      265550 TTTCGTTTGAAGTACCAATTCATATTATAGCTGGAAGAG 265507

```

```

RESULT 15
AX144037 960 bp DNA linear PAT 31-MAY-2001
LOCUS Sequence 2759 from Patent W00134809.
DEFINITION AX144037
ACCESSION AX144037.1 GI:14282890
VERSION AX144037.1
KEYWORDS
SOURCE .
synthetic construct
synthetic construct
ORGANISM artificial sequences.
REFERENCE
1
AUTHORS Kimmberly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 2759 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
Source Location/Qualifiers
1..960
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 336 a 131 c 189 g 304 t
ORIGIN

```

```

Query Match 51.2%; Score 667.6; DB 6; Length 960;
Best Local Similarity 81.2%; Pred. No. 6.8e-106;
Matches 775; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

```

```

QY      352 GTTACAAAGTATGATTAATGAAATGCGTACAGAGCTGATATTTCTATTCATTAGGA 411
Db      1 GTGAATTAAGTTGATATCTTGAATATGCGTAAATGATATCTATTTCTATTTAGGC 60
QY      412 TTTCGTAACCGTATCCAAATATGCGTACATGCTTGTAGTCTTTGGAAGTCTTGTAGAT 471
Db      61 TTTCGTAACCGTATCCAAATATGCGTACATGCTTGTAGTCTTTGGAAGTCTTGTAGAT 120
QY      472 GCAGTTGTTCTCTATTTGTTGGAAGAGAAAGATCTTATGATGAAGATACATTCGA 531
Db      121 GCAGTTGTTGTAAGACCTTAATTAAGAAATCAGAAATCTTATGACGAAGATACATTCAGT 180
QY      532 CTATCATTTATTTGAGACCAAAACGTAGTAATCAAGTTTATGTTAGT 591

```

```

Db      181 CTTTCTATCATCGGTAGACCTAATGTTGTAATCTAGCTTGTCTAATGCTATTTTACGC 240
QY      592 GAACATGCGCTTATCGTTTCTAATGTTTCAGGACCAACGAGACGCTATTGATACAGAG 651
Db      241 GAAGAACCTGTTATTTGTTCTAATGTTGCTGATCAACTGACAGATGCCATTTGATACGAG 300
QY      652 TATAGTTATGATGACAAAGATTAATGTTTAAATCGATACGCTGATGCTGATTAAGAAAGGA 711
Db      301 TACTCTTATGATGACAAAGATTAATGATTAATGATTAATGCTGCGAATGAGAAAAAAGCT 360
QY      712 AAGATATATGATCACTGAGAAATATTCAGTATTTAGAGCTTTAAAGCAATGGAACGT 771
Db      361 AAGTGTATGATGATGACGAGAAAAATATTCGTATTTACGCTTGAATTAAGCAATGGAACGT 420
QY      772 TCAATGTTGTTTATGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 831
Db      421 TCAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 480
QY      832 GTTCAGATATGACATGACAAAGGTTAAACGATGCTGATTTGCTAATTAATGAGGAT 891
Db      481 GTACGTGCTATGACATGACAAAGGTTAAACGATGCTGATTTGCTAATTAATGAGGAT 540
QY      892 ACTGTGAAAAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTC 951
Db      541 ACAGTTGAAAAAGATAGTAAACGATGAAGAAATTTCACTGATGATTTGAAGAAATGAATTT 600
QY      952 CAATTTTATGATTAATGACCAATGCTTTGCTGCTGTAAGAAACGCAAGAAATTAAGT 1011
Db      601 CAATTTTATGATTAATGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY      1012 ACATTAATTCCTTACATTAATGAAGCAAGTGAAGAAACATTAAGAAACGCTTCAAGTTCGA 1071
Db      661 ACATTAATTCCTTATATCAATCAACGCGTGAAGAAATCAATTAAGAAACGCTTCAAGTTCCT 720
QY      1072 ACTTTAATGAAGTCTTACTGATGCAATTTTCATGAACCTTACACCAACGACAAAGCT 1131
Db      721 ACCTTAATGAAGTCTTACTGATGCAATTTTCATGAACCTTACACCAACGACAAAGCT 780
QY      1132 AGAGTTTGAATGCTTTTATGCAACACAACTGCTATAGAACCAACGCAATTTGTGTA 1191
Db      781 AGAAGACTTAATGATTAATGCTTATACCACTGATGCTGATTAAGTGAACCAACGCAATTTGTGTA 840
QY      1192 TTTCGTAATGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1251
Db      841 TTTCGTAATGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 900
QY      1252 CGTGCCGCTTTTGTGTTTGAAGTACACCAATTCATTAATTAAGCTGCAAGAGA 1305
Db      901 CGTAAATGCTTTTGTGTTTGAAGTAAACACCTTATTAATTAATTAATTAATTAATTAATTAAT 954

```

Search completed: September 12, 2003, 19:27:32
 Job time : 5331.98 secs

THIS PAGE BLANK (USPTO)

ACCESSION	AY079759	uncultured marine virus genomic clone SI051pD7L, genomic survey sequence.
VERSION	AY079759.1	GI:24744504
KEYWORDS	GSS,	
SOURCE	uncultured marine virus	
ORGANISM	uncultured marine virus	
REFERENCE	Viruses; environmental samples.	
AUTHORS	1 (bases 1 to 745)	
TITLE	Breibart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M., Mead,D., Azam,F. and Rohwer,F.	
JOURNAL	Genomic analysis of uncultured marine viral communities	
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)	
PUBMED	22294988	
COMMENT	12384570	
Contact:	Rohwer F	
Biologiy Dept.		
San Diego State University		
5500 Campanile Dr, San Diego, CA 92102, USA		
Tel: 6195941336		
Fax: 619595676		
Email: forest@sunstroke.sdsu.edu		
Class: shotgun.		
location/Qualifiers		
1..745		
/organism="uncultured marine virus"		
/mol_type="genomic DNA"		
/db_xref="taxon:186617"		
/clone="SI051pD7L"		
/clone_1lb="Scripts Pier (La Jolla, CA) uncultured virus community"		
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)"		
BASE COUNT	167 a 221 c 226 g 127 t 4 others	
ORIGIN		
Query Match	11.5%; Score 44.4; DB 28; Length 745;	
Best Local Similarity	48.7%; Pred. No. 0.34;	
Matches	114; Conservative 0; Mismatches 120; Indels 0; Gaps 0;	
OY	86 CGGTTACCAAACTCTAATGAATAATCATACAGTCGTGACGATTTCATATTATC 145 	
Db	745 CGGTTCCCGCAAGGCCGAGANAATAGCTTTCAGAGATGCCTTTCCGGCTTCCCTTC 686	
OY	146 TACTTTGTTAACCGCTATATACGACCAGGTTTTTGATTTGATTAATAATTGAGGACCAT 205 	
Db	685 TGCCNTGTTCGCAACGACGACACGGGCGACTTCTGTCGGCGACCGACGCGCAATTTTC 626	
OY	206 TTCAATCGCTTTGTGTCATATCTTTCACGACGCTTAACCATAAATAATTAATACCGCTTC 265 	
Db	625 NTGCTCATCGGTGTGATMGCCAGCACCGCATCATGACGAACAATCCGGCGTGGCATC 566	
OY	266 ATCTATGGCGATTTTCTCCCTGCGCTCTAATTTTGTGTTGGAATGGTGATCACC 319 	
Db	565 GTCAATGCTGACTGCTGGTCTGCTTGGCCGATGCGTCCGGCAGGAGGCGTCATC 512	
RESULT 8		
CNS000D1/c		
LOCUS	1101 bp DNA linear GSS 03-JUN-1999	
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR0J116 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL065414	
VERSION	AL065414.1	GI:4938827

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 1101)	Genoscope.	Direct Submission	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library was prepared by Kazuo Oseguera and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACRPC Resource Center can be found at http://bacrpc.med.buffalo.edu/drosophila_bac.htm .		
FEATURES	Location/Qualifiers		
source	1..1101		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/clone="BACR01J16"		
	/clone_11b="RPCI-98"		
	/note="end : TGT3"		
BASE COUNT	280 a	104 c	123 g
ORIGIN	211 t	383 others	
Query Match	10.7%; Score 41.6; DB 29; Length 1101;		
Best Local Similarity	13.8%; Pred. No. 2.1;		
Matches	34; Conservative	128; Mismatches	84; Indels 0; Gaps 0.
QY	138 ATATTATCTATTGGTATACCGGATGATGACGCGGTTTGTAGATTTGTATGAATTTGA 197		
DB	1099 MTNNHATGTTGNNHNTNTGMMTNNHMTGTTMTGTTNTGNTNNHMTGTTNTGTTTGC 104d		
QY	198 GCGACGATTCATCGGCTTTGTGTCATGCTTCACGCGGATTCATGATGATGATGATGAC 257		
DB	1039 MGNHNNHMMAMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMH 980		
QY	258 TCGGCTTCATGATGCGGATTCGCGGCTTCGCGGCTTCATGATGATGATGATGATGATGAC 317		
DB	979 TMTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 920		
QY	318 CCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 377		
DB	919 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860		
QY	378 GAATGA 383		
DB	859 MHTATG 854		
RESULT 9	A2525216		
LOCUS	A2525216 653 bp DNA linear		
DEFINITION	241PbD05 Pb MBN #21 Plasmidium berghel genomic 3', genomic survey sequence.		
ACCESSION	A2525216		
VERSION	A2525216.1 GI:13965843		
KEYWORDS	GSS.		
SOURCE	Plasmidium berghel		
ORGANISM	Plasmidium berghel		

[illegible]

KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101)
REFERENCE	Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.flytiffy.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPac Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES	location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone_11b="RPc1-98" /clone_11d="RPc1-98" /note="end : T7"
BASE COUNT	303 a 195 c 173 g 280 t 150 others
ORIGIN	
Query Match	10.4%; Score 40.2; DB 29; Length 1101;
Best Local Similarity	28.3%; Pred.No. 5;
Matches	80; Conservative 72; Mismatches 131; Indels 0; Gaps 0;
OY	20 AAAATGGAACAACCTGCATCTAACCACTCAACAGACTTAACCATTGACCCGTGAT 79 :: :: :: :: Db 819 AAACAAMAAHAAAAAAMAAAACMAAMKCCYCCCCCAAATAATTAMMCMCMATWT 878 OY 80 CGGATAGCGGTTCACCAAACTCTAATGAATAGAATCATACACGGTGTGACATTTCAT 139 ::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: Db 879 TTTNNHTTTATTTHAAHAAAAAAAAAAAAAAAAAACCSTRAHMAACAMTWITTYCCAC 938 ::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: OY 140 ATATATCTACTTTTGTTAAACCGCTAAATACGACCGGTTTTTTAGATTGTTAAATTTGAGC 199 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: Db 939 ACNHHMMMAACAHNAVCSTGKTWANHMCAMYTWTMTTMMTTTTMMWWMMHMMMAAAAT 998 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: OY 200 GACCATTTCAATCGCTTTGTGCAATCTTCAACGCAAGTTAAACATAAATAATATACATC 259 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: Db 999 AMATATATTTTTTTTTTTTTTAAATATAGCTHAMAWTCTNHUYWATCTMYACWCACAYA 1058 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: OY 260 CGCTTCATCTAGGCGATTTCTGCCCTGGCCGCTATATTTGATTT 302 ::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: Db 1059 TTMHNYTMTHTATTTTYYTLYHTCATATAYATTAATACCKWCT 1101
RESULT 13	
CNS006TJ	
LOCUS	CNS006TJ 938 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #
DESCRIPTION	BACR1411 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL065906
VERSION	AL065906.1 GI:4944874
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)

This clone is available at CHGC in Shanghai.

FEATURES

source

1..709

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="FHTBAE05"

/tissue_type="hypothalamus"

/dev_stage="Fetal"

/lab_host="BM25.8"

/clone_lib="FHTB"

/note="vector: pT7.1Ex2; Site_1: sf11A; Site_2: sf11B1"

BASE COUNT

185 a 149 c 145 g 227 t 3 others

ORIGIN

Query Match

10.1%; Score 39.2; DB 9; Length 709;

Best Local Similarity 56.1%; Pred.No. 8.3;

Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
QY      218 TGTCAATCCTTCACGACGCTTAACCATATAAATAATACATCCGCTTCATCTATGGCGAT 277
          |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB      247 TGTCAATCTAGTACTATATTTAGTAGTGAACCAATACATCTCTTTCTTATGCTTG 306
          |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY      278 TTCTGCGTGGCGCTTAATTTGTGTTGGAAATGATGCGATCAGCAATTTCAATACCACTGT 337
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB      307 TTTCTCTTTCATGTAAATTTTGTCTGTATGAGACGACCATGATAGAGTTGGCAATTGCA 366
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY      338 ATCATATATATT 349
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB      367 AACATTTTATTT 378
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
```

Search completed: September 12, 2003, 20:43:34

Job time : 1049.84 secs

THIS PAGE BLANK (USPTO)